

کشف دلایل رویگردانی مشتری از خدمات بانکداری با ترکیب روش‌های داده‌کاوی و تحقیق پیمایشی

محمد مهدی سپهری^{۱*}، اشرف نوروزی^۲، بابک تیمورپور^۳، سروناز چوبدار^۴

- ۱- دانشیار بخش مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، ایران
- ۲- دانشجوی دکتری مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، ایران
- ۳- دکتری مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، تهران، ایران
- ۴- کارشناس ارشد مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ایران

دریافت: ۸۹/۳/۲۵

پذیرش: ۹۰/۳/۲۲

چکیده

رویگردانی مشتری به مفهوم خاتمه روابط وی با سازمان و رویکرد وی به سمت سایر تأمین‌کنندگان در نتیجه نارضایتی وی از خدمات سازمان مفهومی است که شناسایی و درک دلایل بروز آن شرط بقای کسب‌وکار در شرایط رقابتی است. تاکنون مطالعات فراوانی در خصوص پیش‌بینی نرخ رویگردانی مشتری صورت گرفته است ولی به شناسایی دلایل بروز این پدیده تاکنون کم‌تر پرداخته شده است. برخی از محققان دلیل این امر را عدم کفایت محتوای داده موجود در پایگاه‌های داده تراکنشی مشتریان برای این منظور دانسته‌اند که لزوم بهره‌گیری از نتایج تحقیقات پیمایشی مناسب برای این منظور را آشکار می‌سازد. در طول تحقیق حاضر تلاش شده است تا با بهره‌گیری از تحقیق پیمایشی پرسشنامه محور نظرات مشتریان در خصوص سطوح رضایت‌مندی و مؤلفه‌های تأثیرگذار بر رویگردانی مشتری ارزیابی شده و از نتایج حاصل برای تحلیل تأثیر عوامل مختلف بر رویگردانی بهره گرفته شود. مشتریان حساب‌های جاری اشخاص بانک کشاورزی به عنوان مورد مطالعاتی این تحقیق انتخاب شده‌اند. رویکرد استفاده شده برای شناسایی علل رویگردانی نیز استخراج قواعد منتهی به رویگردانی در گروه‌های مختلف مشتری بوده است. برای این منظور از کاربرد تکنیک درخت تصمیم با متغیر هدف برچسب رویگردانی استفاده شده است. اعتبارسنجی نتایج حاصل نیز با آزمودن آن بر روی داده‌های اعتبارسنجی و محاسبه نرخ lift و خطای کلی به دست آمده است. قواعد



استخراجی نشان از تمایل به رویگردانی در بخش بزرگی از مشتریان بانک کشاورزی دارد که عمده‌ترین آن (که تأثیر آن به‌ویژه در گروه‌های با درآمد بالاتر بیشتر مشاهده می‌شود) طرز برخورد کارمندان بانک بوده است.

کلیدواژه‌ها: مدیریت رویگردانی مشتری، پیش‌بینی نرخ رویگردانی، شناسایی علل رویگردانی، بانک‌داری، داده‌کاوی.

۱- مقدمه

جذب مشتریان جدید و حفظ و نگهداری مشتریان ارزشمند قدیمی هر دو دارای اهمیت می‌باشند. در شرایطی که شرکت‌های نوظهور بیش‌تر بر روی کسب مشتریان جدید تکیه دارند، شرکت‌های با سابقه‌تر بر حفظ مشتریان فعلی تکیه دارند. واقعیت این است که این مشتریان سود بیش‌تری برای مؤسسه تأمین می‌کنند و اغلب فرصت‌های فروش بیش‌تری در برابر آنان در اختیار می‌باشد [۱، صص ۳۹۴-۳۹۹؛ ۲، صص ۸۳-۹۹؛ ۳، صص ۲۸۷-۳۰۱؛ ۴، صص ۶۲۳-۶۳۳؛ ۵، صص ۳۷۵-۳۸۵؛ ۶، صص ۲۵۹۲-۲۶۰۲؛ ۷، صص ۴۸۳-۵۰۲].

براساس مطالعات متخصصان بازاریابی، بیش‌تر شرکت‌ها به طور میانگین سالیانه در حدود ۲۵ درصد از مشتریان خود را از دست می‌دهند [۸، صص ۲۹۳-۳۰۲]. طی سال‌های متمادی تمرکز اصلی مطالعات وفاداری بر روی تمایل مشتریانی که قصد شروع روابط با سازمان را داشته‌اند و همچنین حفظ و تعمیق این روابط از طریق افزایش رضایت‌مندی بوده است. یافتن و حفظ مشتریان جدید به‌ویژه زمانی که مشتریان به حداکثر تعداد خود می‌رسند، بسیار دشوار و پرهزینه می‌شود. در چنین نقطه‌ای از چرخه عمر یک کسب‌وکار، اولویت بیش‌تر بر حفظ مشتریان قدیمی می‌باشد [۹، صص ۸-۱۱؛ ۱۰، صص ۱۲۵-۱۳۴؛ ۱۱، صص ۳۷۴-۳۹۰؛ ۱۲، صص ۱۴۵-۱۵۹؛ ۱۳، صص ۸۲-۹۷؛ ۱۴، صص ۱۵۱-۱۵۹؛ ۱۵، صص ۵۶۹-۵۹۰؛ ۱۶، صص ۸۰۷-۸۰۷؛ ۱۷، صص ۱۰۳-۱۱۲؛ ۱۸، صص ۵۰-۵۸]. برخی تحقیقات بر این نکته مهم تکیه کرده‌اند که توجه خاصی باید روی مشتریانی صورت گیرد که اگر چه در این لحظه وفادار تلقی شده‌اند و منافع زیادی برای سازمان تأمین می‌کنند ولی به دلیل بروز برخی مشکلات دارای احتمال بالایی از ترک سازمان در آینده خواهند بود [۱۹، صص ۱۱۷-۱۳۱].

هدف این تحقیق شناسایی و تحلیل عمده‌ترین دلایل رویگردانی مشتریان از خدمات حساب‌های جاری بانک کشاورزی به سوی سایر مؤسسات می‌باشد. حوزه حساب‌های جاری بخشی بسیاری مهم از بازار خدمات بانکی را تشکیل می‌دهد که بروز نارضایتی در میان مشتریان این گروه و روی آوردن آن‌ها به سایر مؤسسات، هزینه بسیار سنگینی برای بانک به دنبال دارد. با توجه به این که ممکن است اولویت‌ها و عوامل تأثیرگذار بر تصمیم مشتریان به ادامه روابط برای گروه‌های مختلف مشتری یکسان نباشد، کاربرد تکنیک‌های مرسوم آماری برای تلاش در سنجش تأثیر عوامل مختلف بر تصمیم‌های مشتری چندان راهگشا نمی‌باشد [۲۰، صص ۱۱۰۴۵-۱۱۰۵۶؛ ۲۱، صص ۱۴۲۱۸-۱۴۲۲۴؛ ۲۲، صص ۹۲۷۴-۹۲۸۰].

از این رو در تحقیق حاضر تلاش شده است تا با کمک تکنیک‌های داده‌کاوی نخست اقدام به خوشه‌بندی مشتریان به رده‌های گوناگون شده و در ادامه تأثیر عوامل مختلف بر تصمیم رویگردانی ایشان سنجش شود. با توجه به این که در حال حاضر، بانک بی‌توجه به بخش‌های مختلف مشتری، سیاست یکسانی را در ارتباط با ایشان و اجرای برنامه‌های وفاداری خود اتخاذ می‌کند، ارزش این تحقیق در تحلیل عوامل مؤثر بر رویگردانی مشتریان خوشه‌های مختلف آشکار می‌شود. همچنین تحقیق حاضر اولین نمونه تحقیق با به‌کارگیری توأم روش پیمایشی و داده‌کاوی برای ارزیابی سطح تأثیر عوامل تأثیرگذار بر وفاداری یا رویگردانی مشتری در ایران است که با بهره‌گیری از تکنیک‌های داده‌کاوی بر روی داده‌های حاصل از تحقیق پیمایشی اقدام به بخش‌بندی مشتریان و سپس ارزیابی و تحلیل تأثیر عوامل مختلف بر وفاداری ایشان و مقایسه نتایج حاصل برای خوشه‌های مختلف مشتری شده است.

۲- مبانی نظری و پیشینه تحقیق

۲-۱- مفهوم رویگردانی مشتری

رویگردانی مشتری، عبارتی است که برای بیان از دست رفتن مشتری به علت‌های گوناگون به کار می‌رود. به بیان دقیق‌تر، مشتری رویگردان فردی است که در آستانه ترک سرویس‌های ما و استفاده از سرویس رقبا قرار دارد [۲۵، صص ۲۹۰۲-۲۹۱۷]. رویگردانی را می‌توان به دو دسته اصلی تقسیم‌بندی کرد: رویگردانی ارادی^۱ و رویگردانی اجباری^۲ [۲۵،



صص ۲۹۰۲-۲۹۱۷]. رویگردانی اجباری زمانی رخ می‌دهد که سازمان خود از ارائه سرویس به مشتری به دلایلی نظیر استفاده سوء از خدمات و یا عدم پرداخت صورت حساب‌ها خودداری کند. رویگردانی ارادی زمانی رخ می‌دهد که مشتری تصمیمی آگاهانه برای خاتمه سرویس‌گیری و رویکرد به مؤسسات رقیب اخذ می‌کند. دلایل این رویگردانی می‌تواند مسائلی نظیر تغییرات تکنولوژی، مسائل اقتصادی، فاکتورهای کیفی، نوع و پوشش سرویس‌دهی و یا حتی تجربه‌ای بد در برخورد با کارکنان مرکز تلفنی یک سازمان باشد [صص ۷۵۱-۷۶۵]. در این تحقیق هر جا نام رویگردانی برده می‌شود، منظور رویگردانی ارادی مشتریان ناشی از نارضایتی ایشان از خدمات ارائه شده می‌باشد.

۲-۲- مفاهیم رضایت‌مندی و وفاداری مشتری

به طور کلی رضایت مشتری تحت تأثیر توقع مشتری یا انتظارات وی قبل از دریافت سرویس و ارزش دریافت شده به وسیله وی پس از ارائه خدمت تعیین می‌شود. بنابراین رضایت مشتری را می‌توان به شکل فرمول زیر تعریف کرد:

رضایت مشتری = ارزش درک شده به وسیله وی از خدمت ارائه شده - انتظارات وی [صص ۴۶۰-۴۶۹].

با توجه به این تعریف، چنانچه کالا و خدمات دریافت شده از جانب مشتری انتظارات او را برآورده سازد، در او احساس رضایت ایجاد می‌شود و در صورتی که سطح خدمت و کالا پایین‌تر از سطح انتظارات مشتری باشد، به نارضایتی وی منجر خواهد شد. رضایت مشتری سنجه‌ای معمول برای اندازه‌گیری عملکرد سازمانی به شکل بازخوردی حاصل از این عملکرد است. همچنین این پارامتر پیش‌بینی کننده‌ای مهم از موفقیت بسیاری از سنجه‌های مالی و عملکردی سازمان می‌باشد [صص ۷۱۸-۷۳۹]. بسیاری از تحقیقات قبلی تأییدکننده تأثیر مثبت رضایت مشتری بر وفاداری مشتری و رفتار خرید وی [صص ۹۳-۲۱۸]، سود [صص ۷-۲۴] و جریان‌های نقدی آینده شرکت [صص ۱۱۵-۱۳۰] بوده‌اند و در عین حال تأثیر منفی آن را بر نرخ شکایات و هزینه تراکنش‌های آینده نشان داده‌اند [صص ۱۰۵-۱۱۱]. وفاداری مشتری زمانی که انتظارات وی در رابطه با تجربه خرید کالا یا خدمات برآورده شود، ایجاد شده و در صورت تداوم آن به صورت مراجعات مکرر توسعه پیدا می‌کند. زمانی

که رابطه وفاداری مشتری شکل گرفت، سازمان باید به طور مداوم اقدام به اندازه‌گیری و بررسی آن به منظور تضمین حفظ و نگهداری این رابطه نماید [۱۲، صص ۱۴۵-۱۵۹]. در واقع نگهداری مشتری که به مفهوم حفظ روابط وفادارانه شکل گرفته با مشتریان می‌باشد، قلب CRM و یکی از دلایل اصلی شکل‌گیری این مبحث را تشکیل می‌دهد؛ به عبارتی می‌توان CRM را به شکل تلاشی در جذب مشتریان جدید و پس از آن توسعه و نگهداری وفاداری این مشتریان تعریف نمود [۲۳، صص ۴۷-۵۷]. تعاریف فوق به‌خوبی ارتباط نزدیک دو مفهوم رضایت‌مندی و وفاداری مشتری را نشان می‌دهد تا جایی که در بسیاری موارد از اندازه‌گیری شاخص رضایت‌مندی در پایش وفاداری مشتری استفاده می‌شود [۲۹، صص ۹۳-۲۱۸].

۲-۳- پیشینه تحقیق

از جمله تحقیقات قبلی که به مفهوم احتمال رویکرد مشتری به سمت سایر مؤسسات در صنعت بانکداری در ایران نگاهی داشته‌اند می‌توان به تحقیق امیر شاهی و همکاران (۸۸) اشاره کرد [۲۳، صص ۶۱-۷۶]. ایشان در تحقیق خود با اشاره به رقابتی بودن صنعت بانکداری و احتمال رویکرد مشتریان به سوی رقبای، به بررسی عوامل مؤثر بر افزایش اعتماد مشتریان کلیدی به عنوان راهکاری برای افزایش وفاداری مشتریان پرداخته‌اند. سهرابی و همکاران (۸۹) نیز با بیان نیاز به متمایزسازی برخورد با مشتریان صنعت بانکداری به منظور پیشگیری از رویکرد آنان به سمت رقبای، اقدام به بخش‌بندی مشتریان در صنعت بانکداری بر اساس ارزش مشتری کرده‌اند تا براساس آن اقدام به تدوین استراتژی‌های مناسب برای برخورد با هر دسته از مشتریان گردد [۲۴، صص ۲۲۳-۲۳۹]. با این وجود تحقیقات داخلی که به‌طور مستقیم به پدیده رویگردانی اختصاص داشته باشند، بسیار اندک بوده است و نویسندگان در مرور ادبیات موضوع موردی از آن را مشاهده نکرده‌اند. در بیش‌تر تحقیقاتی که تاکنون در خصوص مدیریت رویگردانی مشتری از سازمان صورت گرفته است، اقدام به پیش‌بینی نرخ رویگردانی با کمک داده‌های تراکنشی موجود از تعاملات مشتری شده و در حوزه شناسایی دلایل شکل‌گیری رویگردانی و تلاش در جهت رفع عوامل ایجادکننده این پدیده مطالعه چندانی صورت پذیرفته است [۲۵، صص ۲۹۰۲-۲۹۱۷]. این در حالیست که مطالعه نرخ رویگردانی مشتری به تنهایی چیزی فراتر از دیدی کلی در خصوص



کسب و کار ارائه نمی‌دهد و هر گونه تلاش در رفع علل ایجادکننده این پدیده منوط به درک دلایل و شرایط بروز پدیده روگردانی است [۸، صص ۲۹۳-۳۰۲]. تحقیقات درخصوص روگردانی مشتری در اصل به دو دسته اصلی پیش‌بینی نرخ روگردانی و شناسایی علل این پدیده تقسیم می‌شوند: تحقیقات موجود در زمینه پیش‌بینی نرخ روگردانی مشتری و افزایش دقت آن به‌طور عمده به پیش‌بینی نرخ روگردانی براساس سوابق رفتارهای مشتری اختصاص دارند؛ به عنوان مثال داتا و همکاران (۲۰۰۱) طی تحقیق خود به منظور پیش‌بینی نرخ روگردانی در صنعت ارتباطات با بهره‌گیری از داده‌های تراکنشی مربوط به پرداخت‌ها و سوابق مصرف مشتری و با کلاس‌بندی مشتریان با روش درخت تصمیم اقدام به پیش‌بینی نرخ روگردانی کردند [۳۴، صص ۴۸۵-۵۰۲]. بخش عمده‌ای از تحقیقات صورت گرفته در زمینه پیش‌بینی روگردانی مشتری نیز به تلاش در جهت افزایش دقت الگوهای پیش‌بینی اختصاص دارد؛ به عنوان مثال کوسمنت و ون دن پل (۲۰۰۸) برای شناسایی مشتریان روگردان تلاش کرده‌اند تا با اضافه کردن اطلاعات ایمیل‌های مرکز تماس به سیستم سنتی پیش‌بینی کننده نرخ روگردانی مشتری، دقت این الگو را افزایش دهند [۳۵، صص ۱۶۴-۱۷۴]. گلادی و همکاران (۲۰۰۸) نیز در تحقیق خود رویکرد جدیدی را به کار گرفتند و طی آن ارزش دوره زمانی عمر مشتری را یکی از پارامترهای دخیل در پیش‌بینی نرخ روگردانی مشتری فرض کردند. طی این دیدگاه کاهش ارزش مشتری به عنوان معیار روگردانی فرض می‌شود [۳۶، صص ۴۰۲-۴۱۱]. بخشی از تحقیقات این حوزه نیز تمرکز بر توسعه تکنیک‌های جدید پیش‌بینی کننده داشته‌اند؛ به عنوان مثال وربکه و همکاران (۲۰۱۱) برای شناسایی مشتریان روگردان از ترکیبی از تکنیک‌های مختلف استنتاج قواعد استفاده کردند و کارایی آن را با سایر روش‌های معمول داده‌کاوی در این زمینه مورد مقایسه قرار دادند [۳۷، صص ۲۳۵-۲۳۶]. بوک و ون دن پل (۲۰۱۱) نیز در تحقیق خود از ترکیبی از روش‌های دسته‌بندی (Rot Boost و Rotation forest) در زمینه پیش‌بینی روگردانی مشتری استفاده کردند و کاربرد تکنیک جدید را با تکنیک‌های مرسوم در این زمینه مقایسه نمودند [۳۸، صص ۱۲۲۹۳-۱۲۳۰۱]. در مورد زمینه دوم که تحقیق حاضر نیز به آن اختصاص دارد؛ یعنی شناسایی علل به وجود آورنده روگردانی، تاکنون تحقیقات بسیار اندکی صورت گرفته است، نمونه ای از این نوع تحقیقات، مطالعه شین و کیم (۲۰۰۸) است که طی آن به منظور بررسی

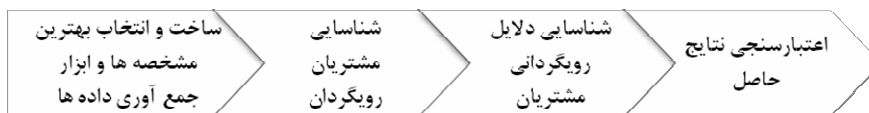
تأثیر عوامل مختلف بر رویگردانی مشتری، با فرض تأثیر عواملی مانند سطوح رضایت، موانع تغییر تأمین‌کننده و مشخصه‌های دموگرافیک بر رویگردانی، سطح این تأثیر سنجیده شده است. برای این منظور با جمع‌آوری اطلاعات پرسشنامه به شکل تلفنی از میان ۵۲۰ نمونه، تلاش شده است تا تأثیر متغیرهای مستقل فوق بر متغیر وابسته صفر و یک تمایل مشتری به ترک سازمان در آینده سنجش شود [۳۹، صص ۸۵۴-۸۷۴]. کیم و یون (۲۰۰۴) نیز در طول تحقیق خود به منظور شناسایی عوامل تأثیرگذار بر رویگردانی در صنعت ارتباطات، از آزمون کای مربع برای آزمودن تأثیر مؤلفه‌های کیفیت خدمت و مشخصه‌های دموگرافیک مشتریان بر تصمیم رویگردانی ایشان استفاده کرد. نویسندگان اظهار داشته‌اند که هیچ روش مرسوم برای شناسایی دلایل رویگردانی مشتری وجود نداشته است و آن‌ها نیز تلاش کردند تا از روشی تجربی برای شناسایی این علل و اندازه‌گیری میزان تأثیر آن‌ها بهره‌گیرند [۲۶، صص ۷۵۱-۷۶۵]. نمونه‌ای دیگر از این مطالعات، تحقیق چیانگ و همکاران (۲۰۰۴) برای شناسایی دلایل رویگردانی مشتری در حوزه بانکداری اینترنتی است که طی آن از کاربرد روش قواعد انجمنی بر داده‌های تراکنشی مشتریان برای کشف عمده‌ترین الگوهای متوالی منجر به رویگردانی استفاده شده است.

براساس نتایج حاصل، عمده‌ترین دلیل رویگردانی موضوع اشتباه وارد کردن رمز عبور^۲ برای چند بار بوده است که به دلیل ترس از دست دادن رمز عبور تکرار نشده است. با این وجود، تحقیق مذکور تنها موفق بوده است تا دلایلی را که براساس لاگ فایل‌های اینترنتی کاربران نشانه‌های آن قابل مشاهده می‌باشد، مشخص سازد. به بیان دیگر کاربران دیگری که به دلایلی دیگر نظیر نارضایتی از خدمات الکترونیک، بی‌اعتمادی به آن و یا هر عامل دیگر علاقه‌ای به این سیستم ندارند، در این تحقیق ناشناخته باقی مانده‌اند [۸، صص ۲۹۳-۳۰۲]. هادن و همکاران (۲۰۰۵) در مقاله مروری خود بر تکنیک‌ها و الگوهای استفاده شده در ادبیات مدیریت رویگردانی مشتری، یکی از دلایل کم‌تر پرداخته شدن به موضوع شناسایی دلایل رویگردانی مشتری را عدم کفایت محتوای داده موجود در پایگاه‌های داده تراکنشی مشتریان برای بیان دلایل رویگردانی مشتری دانسته و این موضوع را یکی از زمینه‌های آتی



تحقیقاتی معرفی کرده‌اند که نیاز به بهره‌گیری از نتایج تحقیقات پیمایشی مناسب برای استخراج دلایل این پدیده دارد [۲۵، صص ۲۹۰۲-۲۹۱۷].

نویسندگان اظهار کرده‌اند که شناسایی مشتریان رویگردان، خود پیش زمینه‌ای برای مطالعه شناسایی دلایل رویگردانی مشتری است که طی آن باید تأثیر عوامل مختلف بر نرخ رویگردانی مشتریان مورد مطالعه قرار گیرد. چارچوب توسعه داده شده در تحقیق ایشان برای مدیریت رویگردانی مشتری فرایندی پنج مرحله‌ای است که شامل شناسایی بهترین داده مورد استفاده^۶، شناخت معنایی داده^۷، انتخاب مشخصه‌ها^۸، توسعه الگوی پیش‌بینی کننده^۹ و اعتبارسنجی نتایج حاصل^{۱۰} است. در طول تحقیق حاضر، این چارچوب اولیه مورد استفاده قرار گرفته است. با این وجود به دلیل بهره‌گیری از روش تحقیق پیمایشی پرسشنامه محور برای جمع‌آوری داده‌های مورد نیاز، سه مرحله اول چارچوب مورد استفاده در قالب گام "شناسایی بهترین داده و ساخت و انتخاب بهترین مشخصه‌ها" در نظر گرفته شده است. همچنین در این‌جا توسعه الگوی پیش‌بینی کننده در دو مرحله متوالی تخمین نرخ رویگردانی مشتریان (شناسایی مشتریان رویگردان) و توسعه الگوی شناسایی کننده دلایل این رویگردانی صورت پذیرفته است. در شکل ۱ گام‌های پیموده شده در طول این تحقیق براساس چارچوب فوق نمایش داده شده است.



شکل ۱ چارچوب پیشنهادی به منظور شناسایی دلایل رویگردانی مشتریان

جزئیات مرتبط با اجرای این چهار مرحله در ادامه و طی بخش‌های ۴، ۵-۱ و ۵-۳ و ۶ به تفصیل بیان شده است.

۳- روش تحقیق

مقاله حاضر همان گونه که در بخش‌های پیش به آن اشاره شد، از ترکیبی از روش‌های پیمایشی پرسشنامه محور و داده‌کاوی برای دستیابی به اهداف خود بهره می‌برد. در تحقیق

پیمایشی پرسشنامه محور، اقدام به جمع‌آوری داده در قالب مجموعه‌ای از سؤالات ساختار یافته از پاسخ‌گویان می‌شود [۴۰، صص ۳۱۲-۳۲۰]. برای تحلیل داده‌های به دست آمده از پرسشنامه‌های فوق، به گونه‌ای که بتوان عوامل گوناگون مؤثر بر تصمیم وفاداری یا رویگردانی مشتریان رده‌های گوناگون را ارزیابی کرد، نیز از تکنیک‌های داده‌کاوی مرتبط استفاده شده است. داده‌کاوی را به صورت کلی می‌توان به عنوان یک قابلیت جستجوی پیچیده داده که از الگوریتم‌های آماری برای کشف روابط مخفی بین داده‌ها استفاده می‌کند، تعریف کرد. داده‌کاوی الگوها و روابط نهفته در میان داده‌ها را آشکار می‌سازد و در واقع بخشی از یک فرایند بزرگ‌تر با نام کشف دانش می‌باشد که کل گام‌های لازم برای به دست آوردن نتایج معنادار از داده‌ها را شامل می‌شود [۷، صص ۴۸۳-۵۰۲]. در این تحقیق از دو تکنیک داده‌کاوی k - میانگین برای خوشه‌بندی مشتریان به خوشه‌های مختلف دموگرافیک و درخت تصمیم برای تحلیل سطح تأثیر رضایت‌مندی از مؤلفه‌های گوناگون بر نرخ وفاداری مشتری استفاده شده است. همچنین تحقیق حاضر به دلیل این که به بررسی ویژگی‌های مشتریان حساب‌های جاری بانک کشاورزی و عوامل مؤثر بر تصمیم ایشان به رویگردانی می‌پردازد، می‌تواند به مسئولان بانک جهت اخذ تصمیم‌های مرتبط و اجرای برنامه‌های وفاداری متناسب با گروه‌های مختلف مشتریان کمک کند، در مجموعه پژوهش‌های کاربردی قرار می‌گیرد. جامعه آماری این تحقیق دارندگان حساب‌های جاری اشخاص فعال بانک کشاورزی در سطح شهر تهران و ۱۰ شعبه تجاری دارای بالاترین نرخ تعاملات بانکی فوق (با پیشنهاد مرکز تحقیق و توسعه بانک در انتخاب شعب) بوده است که با انتخاب نمونه‌ها به روش تصادفی ساده در مجموع ۲۵۰ پرسشنامه به دست آمده است. همچنین برای آزمون روایی سؤال‌ها از روش اعتبار محتوا استفاده شده است. برای سنجش اعتبار محتوای پرسشنامه نیز از نظرات ۱۶ نفر از متخصصان، مدیران و کارشناسان خبره امور بانکی استفاده شده است و با انجام مصاحبه‌های مختلف و کسب نظرات افراد یادشده، اصلاحات لازم به عمل آمده است. به این ترتیب روایی پرسشنامه فوق با ارزیابی صاحب‌نظران و متخصصان امور بانکی تأیید و از روش آلفای کرونباخ برای تعیین پایایی آن استفاده شده است که عدد به دست آمده (۳۷/۷٪) پایایی آن را تأیید می‌کند [۴۱]. مشخصه‌هایی که در



پرسشنامه نهایی داده‌های مربوط به آن‌ها جمع‌آوری شده است، به شرح مندرج در جدول ۱ می‌باشد. در این صورت به منظور در دست داشتن داده‌های آزمون به منظور اعتبارسنجی نتایج حاصل از این مجموعه، اقدام به توزیع ۱۰۰ برگه پرسشنامه در دو شعبه بانک شده است که با نرخ بازگشت ۸۴ درصد، از نتایج حاصل به منظور اعتبارسنجی الگوی پیشنهادی پس از استخراج قواعد رویگردانی استفاده شده است.

جدول ۱ جزئیات مشخصه‌های مورد اندازه‌گیری با کمک پرسشنامه

نام مشخصه	مقادیر مشخصه
مشخصه‌های اجتماعی - دموگرافیک مشتریان	
جنسیت	مرد/ زن
سن	بین ۲۰ تا ۳۰، بین ۳۰ تا ۴۰، بین ۴۰ تا ۵۰، بین ۵۰ تا ۶۰ و بالاتر از ۶۰
شغل	کشاورز و دامپرور، کاسب و بازاری، مدیر، پزشک و مهندس، فرهنگی، دانشجو، کارمند، کارگر، بیکار، سایر
تحصیلات	زیر دیپلم، دیپلم، فوق دیپلم، لیسانس، فوق لیسانس، دکتری
متوسط سطح درآمد ماهانه	کم‌تر از ۳۰۰ هزار، بین ۳۰۰ تا ۶۰۰ هزار، بین ۶۰۰ هزار تا ۱ میلیون، بین ۱ تا ۳ میلیون، بالاتر از ۳ میلیون تومان
منطقه زندگی	شمال، غرب، مرکز، شرق، جنوب تهران، خارج از تهران
مشخصه‌های نشانگر بروز رویگردانی	
تمایل به توسعه فعالیت بانکی در طول ۵ سال آینده	بانک کشاورزی، سایر بانک‌های دولتی، بانک‌های خصوصی/ سایر مؤسسات مالی و اعتباری
میزان توصیه بانک کشاورزی به دیگران	خیلی زیاد، زیاد، تاحدودی، نسبتاً کم، کم، اصلاً
دارا بودن حساب جاری نزد سایر بانک‌ها	بله/خیر
سطح رضایت از حساب جاری نزد سایر بانک‌ها*	
مشخصه‌های بیانگر دلایل وقوع رویگردانی	
عمده‌ترین مشکل مشاهده شده**	مشکلات مرتبط با اخذ دسته چک/ اخذ وام/ سرعت پایین عملیات/ اشتباهات نیروی انسانی/ خرابی‌های سیستم
بیشترین فعالیت بانکی در حال حاضر	بانک کشاورزی/ سایر بانک‌های دولتی/ بانک‌های خصوصی/ سایر مؤسسات مالی و اعتباری
طرز رفتار و نحوه برخورد کارکنان شعب با مشتریان بانک*	
جواب‌گویی مناسب به پرسش‌ها / مشکلات مشتریان*	
میزان سود سپرده‌ها*	

ادامه جدول ۱

نام مشخصه	مقادیر مشخصه
نحوه اعطای تسهیلات و سود آن*	
سرعت در انجام فعالیت‌های بانکی*	
کیفیت ارائه خدمات (دقت و صحت انجام عملیات بانکی)*	
موقعیت مکانی و دسترسی به بانک و خدمات آن*	

* داده‌های این فیلدها دارای مقیاس لیکرت ده گزینه‌ای از کاملاً ناراضی تا کاملاً راضی هستند.
** این فیلد به شکل سؤال باز متنی طراحی شده و پس از تکمیل پرسشنامه‌ها در پنج گروه دسته‌بندی شده است.

۴- شناسایی بهترین داده و ساخت و انتخاب مشخصه‌ها

همان گونه که بیان شد برای شناسایی دلایل رویگردانی مشتری، نخست باید مشتریان رویگردان را از غیر آن شناسایی کرد تا بتوان علل تأثیرگذار بر آن را مورد سنجش و ارزیابی قرار داد. در جدول ۲ شیوه متمایز ساختن مشتری رویگردان از سایر مشتریان در تعدادی از تحقیقات کلیدی معرفی شده است.

جدول ۲ شیوه تمایز مشتری رویگردان از سایر مشتریان در ادبیات

شین و کیم (۲۰۰۸) [صص ۳۹-۸۵۴-۸۷۴] اسپیلا و همکاران (۲۰۰۷) [صص ۴۲-۵۱۳-۵۲۳] کیم و یون (۲۰۰۴) [صص ۲۶-۷۵۱-۷۶۵] لیو (۲۰۰۹) [صص ۴۳۷-۴۳۸۱-۴۳۸]	پاسخ صریح مشتری به ترک سازمان در آینده	شیوه تمایز مشتری رویگردان از سایر مشتریان
کوسمنت و ون دن پل (۲۰۰۸) [صص ۱۶۴-۱۷۴-۱۷۴] بورز و ون دن پل (۲۰۰۸) [صص ۴۴-۵۱۴-۵۱۴] وانگ و همکاران (۲۰۰۹) [صص ۱۶-۸۰۷۱-۸۰۷۱-۸۰۷۵]	عدم تمدید قرارداد در پایان دوره آن	



ادامه جدول ۲

چو و همکاران (۲۰۰۵) [صص ۴۵، ۳۵۹- [۳۶۹ داتا و همکاران (۲۰۰۱) [صص ۴۸۵- [۵۰۲ ژی و همکاران (۲۰۰۹) [صص ۴۶، ۵۴۴- [۵۴۴۹ چیانگ و همکاران (۲۰۰۳) [صص ۸، ۲۹۳- [۳۰۲	عدم مراجعه به سازمان در طول بازه زمانی مشخص	شیوه تمایز مشتری رویگردان از سایر مشتریان
کیم و یون (۲۰۰۴) [صص ۲۶، ۷۵۱-۷۶۵]	عدم تمایل به توصیه سازمان به دیگران	
گلادی و همکاران (۲۰۰۸) [صص ۳۶، ۴۰۲- [۴۱۱	کاهش ارزش دوره زمانی عمر مشتری	
شین و کیم (۲۰۰۸) [صص ۳۹، ۸۵۴-۸۷۴]	سهولت انجام رویگردانی برای مشتری	

همان گونه که مشاهده می‌شود در روابط به صورت قراردادی، عدم تمدید قرارداد در پایان دوره آن به سادگی به عنوان رویگردانی مشتری معرفی می‌شود. در روابط به صورت غیر قراردادی این موضوع تا حدودی پیچیده‌تر می‌باشد. در این تحقیق با بهره‌گیری از ادبیات موصوف، پاسخ مشتری به سؤال‌های مرتبط با اعلام نظر صریح وی درخصوص ادامه روابط و میزان توصیه بانک به دیگران در کنار دو مشخصه دیگر که امکان‌پذیری وقوع رویگردانی به سمت سایر مؤسسات را نشان می‌دهند؛ یعنی دارا بودن حساب جاری نزد سایر بانک‌ها و سطح رضایت از خدمات حساب جاری نزد سایر بانک‌ها به عنوان مشخصه‌های متمایزکننده مشتری رویگردان از سایر مشتریان مورد توجه قرار گرفته است. همچنین در مرور ادبیات موضوع در این حوزه، مشخصه‌های دموگرافیک مشتریان در بسیاری موارد در توسعه الگوهای مختلف پیش‌بینی کننده نرخ رویگردانی مشتریان مورد استفاده قرار گرفته‌اند (مطالعات [صص ۱۶، ۸۰۷۱؛ ۲۶، ۷۶۵-۷۶۵؛ ۳۵، ۱۶۴-۱۷۴؛ ۳۹، ۸۵۴-۹۷۴؛ ۴۳، ۴۳۷۴-۴۳۸۱؛ ۴۴، ۴۹۷-۵۱۴؛ ۴۵، ۳۵۹-۳۶۹؛ ۴۶، ۵۴۴-۵۴۴۹] نمونه‌هایی از آن هستند).

نتایج حاصل از این مطالعات بیانگر آن است که مشخصه‌های دموگرافیک به میزان زیادی بر وقوع رویگردانی مشتری تأثیرگذار می‌باشد. از این رو در تحقیق حاضر نخست از مشخصه‌های دموگرافیک برای خوشه‌بندی مشتری به گروه‌های مختلف دموگرافیک استفاده شده و در ادامه با محاسبه نرخ رویگردانی مشتری برحسب مشخصه‌های مستخرج از ادبیات که در بخش فوق مورد بحث قرار گرفت، بخش‌بندی اولیه‌ای از بازار با نرخ رویگردانی متمایز برای خوشه‌های مختلف به دست آمده است. در ادامه تلاش شده است تا تأثیر عوامل مختلف بر نرخ رویگردانی مشتریان هر خوشه سنجش شود. در جدول ذیل تعدادی از تحقیقات مختلف که به ارزیابی تأثیر عوامل مختلف بر رویگردانی مشتری پرداخته‌اند، ذکر شده است. همان گونه که مشاهده می‌شود مؤلفه‌های سطوح رضایت مشتری، هزینه‌ها و موانع تعویض تأمین‌کننده به عنوان دو عامل اصلی در این تحقیقات مورد توجه قرار داشته است.

در این تحقیق برای تعیین مؤلفه‌های سطوح رضایت مشتری از نتایج تحقیق جامعی که بر روی نمونه‌ای شامل ۲۷۴۳ نفر از مشتریان بانک کشاورزی در ۷ استان کشور انجام شده، استفاده شده است [۴۷، صص ۱-۳۰]. ایشان در تحقیق خود با کمک ابزار پرسشنامه، مؤلفه‌های رضایت‌مندی مشتری را شناسایی کرده و با کمک تحلیل عاملی ۵۰ متغیر حاصل را به ۷ مؤلفه اصلی تقلیل دادند. این مؤلفه‌ها به عنوان مؤلفه‌های رضایت در تحقیق حاضر مورد استفاده قرار گرفته است. درخصوص هزینه‌ها و موانع تعویض تأمین‌کننده نیز سؤالی صریح در خصوص بانک محل بیش‌ترین فعالیت بانکی و دشواری تغییر آن طی دوره آتی از مشتریان در پرسشنامه سؤال شده است.

جدول ۳ مشخصه‌های استفاده شده در بررسی دلایل وقوع رویگردانی مشتری

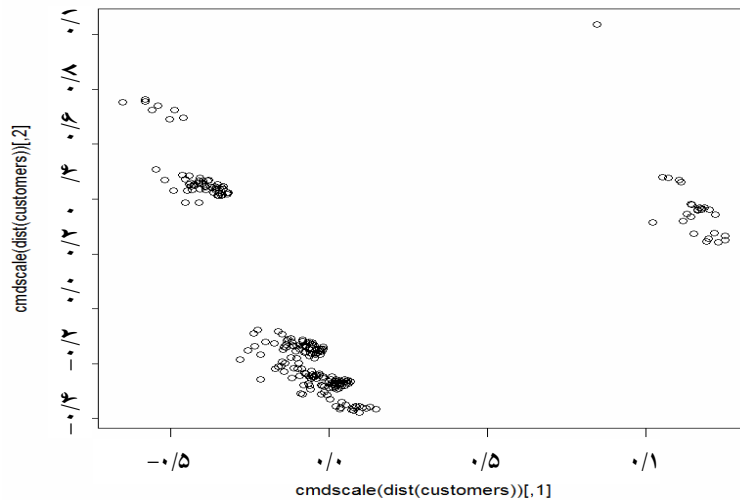
شین و کیم (۲۰۰۸) [صص ۸۵۴-۸۷۴] کیم و یون (۲۰۰۴) [صص ۷۵۱-۷۶۵] لیو (۲۰۰۹) [صص ۴۳۷-۴۳۸]	مؤلفه‌های سطوح رضایت مشتری	مشخصه‌های استفاده شده در بررسی دلایل وقوع رویگردانی مشتری
شین و کیم (۲۰۰۸) [صص ۸۵۴-۸۷۴] اسپیگلر و همکاران (۲۰۰۷) [صص ۵۱۳-۵۲۳] کیم و یون (۲۰۰۴) [صص ۷۵۱-۷۶۵]	هزینه و موانع تعویض تأمین‌کننده	



۵- تجزیه و تحلیل داده‌ها

در این تحقیق از دو نرم‌افزار داده‌کاوی SPSS Clementine 11.1 و R 2.7.1 برای انجام کارهای داده‌کاوی نظیر خوشه‌بندی، ساخت درخت تصمیم و اعتبارسنجی نتایج بهره گرفته شده است.

با توجه به اجرای این تحقیق با کمک پرسشنامه و طراحی فیلدهای داده‌ای به وسیله محقق نیاز چندانی به کارهای آماده‌سازی نبوده است. از این رو تنها دو عمل نرمال‌سازی داده‌ها با روش نرمال‌سازی حداقل-حداکثر (به منظور هم مقیاس کردن داده‌ها) و نیز مصورسازی داده‌ها با کمک روش مقیاس‌گذاری چند بعدی^۹ (به منظور نمایش داده‌ها در فضای دو بعدی) صورت گرفته است. حاصل اجرای این روش برای داده‌های نرمال شده از نظرات مشتریان به شکل ذیل می‌باشد:



شکل ۲ اجرای روش MDS بر روی داده‌های به دست آمده از مشتریان

مبنای محاسبه فواصل بین عناصر در شکل بالا فاصله اقلیدسی آن بوده است.

۵-۱- شناسایی مشتریان رویگردان

در این بخش نحوه محاسبه برچسب رویگردانی برای مشتریان براساس نتایج حاصل از پرسشنامه بیان شده و نحوه خوشه‌بندی این مشتریان به گروه‌های مختلف دموگرافیک معرفی شده است. در بخش آینده براساس برچسب فوق به شناسایی قواعد رویگردانی برای مشتریان هر خوشه پرداخته می‌شود.

۵-۱-۱- ساخت برچسب رویگردانی مشتری

برچسب رویگردانی مشتری براساس مشخصه‌های نشانگر وقوع رویگردانی، محاسبه شده است. در ایجاد این برچسب از متغیرهای زیر استفاده شده است:

X_1 : دارا بودن حساب جاری در سایر بانک‌ها (متغیر صفر و یک- در صورت داشتن حساب جاری در هر بانک دیگر یک و در غیر این صورت صفر)؛
 X_2 : سطح رضایت از حساب جاری در سایر بانک‌ها (بالاترین نرخ رضایت مشتری از سایر بانک‌هایی است که در آن دارای حساب جاری می‌باشد)؛
 X_3 : تمایل به توسعه فعالیت بانکی در آینده (متغیر صفر و یک- در صورت تمایل مشتری به توسعه فعالیت خود در هر بانک دیگر دارای مقدار یک و در غیر این صورت دارای مقدار صفر است)؛
 X_4 : میزان توصیه بانک کشاورزی به دیگران .

مقدار برچسب رویگردانی به سادگی و از مجموع این سه پارامتر (رضایت از سایر بانک‌ها، تمایل به توسعه فعالیت در آینده و میزان توصیه به دیگران) محاسبه شده است. پارامتر صفر و یک دارا بودن حساب جاری در سایر بانک‌ها برای جلوگیری از صفر شدن امتیاز سطح رضایت از سایر بانک‌ها در صورت نداشتن حساب در بانکی دیگر، به این مجموعه افزوده شده است. برچسب رویگردانی بر این اساس به صورت $X_1 \times X_2 + X_3 + X_4$ محاسبه می‌شود که اقدام به نرمال‌سازی آن در محدوده صفر تا یک شده است. براساس این شاخص مشتریان دارای نرخ شاخص بالای ۰/۵ به عنوان مشتریان رویگردان و مشتریان با نرخ شاخص کم‌تر از ۰/۵ به عنوان مشتریان غیر رویگردان فرض شده‌اند. متغیر باینری ساخته شده به این روش به عنوان برچسب رویگردانی به مجموعه داده افزوده شده است. از این برچسب به عنوان متغیر هدف مورد پیش‌بینی در دسته‌بندی مشتریان استفاده شده است.



۵-۲- خوشه‌بندی مشتریان

با توجه به شکل نسبتاً گرد خوشه‌های مشاهده شده در مصورسازی داده‌های مشتریان، روش خوشه‌بندی k - میانگین برای آن انتخاب شده است. همان گونه که در بخش ۳ به آن اشاره شد، در این جا از مشخصه‌های دموگرافیک برای خوشه‌بندی اولیه مشتریان استفاده شده است. روش k - میانگین برای اجرای خوشه‌بندی نیاز به دانستن تعداد خوشه‌ها دارد. از این رو در اجرای این تحقیق تعیین بهینه تعداد خوشه‌های مورد نیاز و اعتبارسنجی تکنیک خوشه‌بندی با کمک معیار سیلوئت انجام شده است. این معیار که یکی از معیارهای متداول اعتبارسنجی خوشه‌بندی است و هر دو معیار فواصل درون خوشه‌ای و برون خوشه‌ای را همزمان در نظر گرفته‌اند و تلاش می‌کنند به‌طور همزمان فواصل درون خوشه‌ای را کمینه و فواصل برون خوشه‌ای را بیشینه سازند [۴۸]. مقدار این ضریب برای هر عنصر از رابطه $s_i = \frac{(b_i - a_i)}{\max(a_i, b_i)}$ محاسبه می‌شود که در آن b_i کم‌ترین فاصله بین خوشه‌ای آن عنصر از بین میانگین فواصل آن از اعضای خوشه‌های دیگر و a_i میانگین فاصله آن عنصر با عناصر دیگر موجود در خوشه خودش است. نتیجه اجرای خوشه‌بندی‌های متوالی برای داده‌های تحقیق و محاسبه معیار سیلوئت در هر مورد به شرح ذیل بوده است.

جدول ۴ محاسبه معیار سیلوئت برای تعداد مختلف خوشه‌ها در روش k - میانگین

تعداد اعضای هر خوشه	عرض سیلوئت برای هر خوشه	نرخ کلی پارامتر سیلوئت
۵۷	۰/۵۷۰	۰/۴۲۲۸
۱۹۳	۰/۳۷۹	
۱۶۹	۰/۵۲۱	۰/۵۵۲۰
۵۵	۰/۵۵۷	
۲۶	۰/۷۳۵	
۲۱	۰/۷۲۷	۰/۵۵۵۷
۱۲۳	۰/۶۳۸	
۵۶	۰/۵۶۲	

ادامه جدول ۴

تعداد اعضای هر خوشه	عرض سیلوئت برای هر خوشه	نرخ کلی پارامتر سیلوئت
۵۰	۰/۲۷۲	۰/۴۶۹۲۵۹
۲۶	۰/۷۲۶	
۵۲	۰/۳۵۴	
۲۰	۰/۲۷۰	
۵۵	۰/۵۱۴	
۹۷	۰/۴۷۷	

به این ترتیب ترتیب معیار سیلوئت خوشه‌بندی با تعداد ۴ خوشه را پیشنهاد می‌کند. خوشه‌های به دست آمده با این روش در طول توضیحات ذیل توصیف شده‌اند:

✓ گروه اول (۱۲۳ عضو) افراد دارای شغل بازاری/ کاسب با بالاترین میانگین سنی و تحصیلات دیپلم و پایین‌تر از آن و سطح درآمد بالا هستند؛

✓ گروه دوم (۵۶ عضو) افراد تحصیلکرده با بالاترین سطح درآمد و مشاغل مهندسی، پزشکی و مدیریتی هستند؛

✓ گروه سوم (۵۰ عضو) مشتریان دارای شغل کارمند و دارای درآمد بین ۳۰۰ تا ۶۰۰ هزار تومان هستند که بیشتر آنان تحصیلات دیپلم دارند؛

✓ گروه چهارم (۲۱ عضو) مشتریان با جنسیت زن و تحصیلات لیسانس، شغل کارمند و آموزشی و دارای پایین‌ترین رده درآمدی هستند.

۵-۳- شناسایی دلایل رویگردانی مشتری

در این تحقیق برای شناسایی دلایل رویگردانی مشتری از قواعد دسته‌بندی منجر به رویگردانی مشتری استفاده شده است. متغیر هدفی که در این دسته‌بندی باید پیش‌بینی شود، برچسب رویگردانی مشتری است. الگوریتم‌های کاندید برای ساخت الگوی دسته‌بندی شامل انواع درخت تصمیم (C5.0، CART، Quest و CHAID)، لجستیک رگرسیون، شبکه‌های عصبی و فهرست تصمیم هستند. معیارهایی که برای این ارزیابی و مقایسه



روش‌های مختلف به کار رفته‌اند، عبارتند از دقت کلی^{۱۱} (درصد رکوردهایی که به‌درستی پیش‌بینی شده‌اند به کل رکوردها)، فضای زیر منحنی ROC، سود^{۱۱} (تابع تجمعی از نرخ اطمینان در مسائل پیش‌بینی) و Lift 20% (نسبت پارامتر مورد پیش‌بینی در ۲۰ درصد اول نمونه‌های با نرخ اطمینان بالا به کل جامعه). در این دسته‌بندی با استفاده از مؤلفه‌های بیانگر دلایل وقوع رویگردانی به عنوان متغیر تصمیم، برچسب رویگردانی وی به عنوان متغیر هدف پیش‌بینی می‌شود. نتیجه مقایسه روش‌های فوق برای این منظور به شرح ذیل می‌باشد:

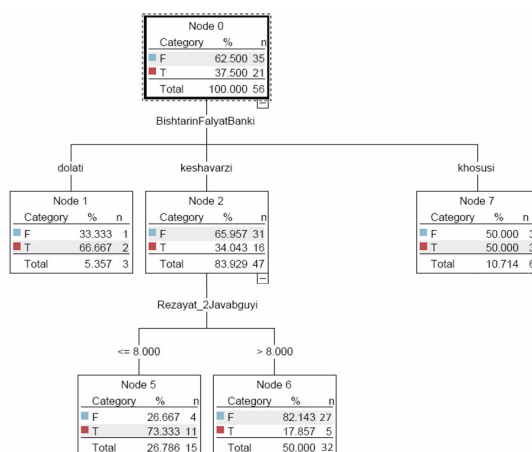
جدول ۵ مقایسه روش‌های مختلف دسته‌بندی برای کاربرد بر داده‌های آموزش

روش	حداکثر سود	Lift 20%	دقت کلی	فضای زیر منحنی ROC
C5.0	۱۹۳/۳۲۳	۲/۱۴	۸۰/۷۲۳	۰/۷۹۱
CART	۱۸۷/۳۶۸	۲/۰۹۵	۸۰/۳۲۱	۰/۷۸۷
شبکه‌های عصبی	۱۱۰	۱/۸۱۸	۷۲/۲۸۹	۰/۷۳۴
QUEST	۱۰۴/۲۵	۱/۷۷۶	۷۳/۴۹۴	۰/۷۴
CHAID	۸۴/۶۴۳	۱/۷۳۴	۷۱/۸۸۸	۰/۷۳۶
فهرست تصمیم	۳۸/۹۴۷	۱/۵۸۲	۶۸/۲۷۳	۰/۶۳۶
لجستیک رگرسیون	۱۰	۱/۰۴۳	۱/۶۰۶	۰/۵۱۸

براساس این مقایسه از روش C5.0 برای دسته‌بندی مشتریان استفاده شده است. در این تحقیق برای افزایش کارایی روش C5.0 از بوس‌تینگ با تکرار ۵ بار استفاده شده است. همچنین به دلیل اهمیت تشخیص مشتریان رویگردان ضریب خطای منفی اشتباه ۱/۵ در نظر گرفته شده است. این ضریب برای خطای مثبت اشتباه برابر یک است (با فرض رویگردان بودن به عنوان مثبت). با اجرای روش دسته‌بندی C5.0 بر روی داده‌های آموزش، چهار درخت تصمیم به‌طور مستقل بر روی هر یک از خوشه‌های مشتریان ساخته شده است. در این الگوریتم از مفهوم بهره اطلاعاتی^{۱۲} یا کاهش آنتروپی^{۱۳} برای تعیین مجموعه بهینه استفاده می‌شود. آنتروپی در یک شکست فرضی از درخت را می‌توان به‌وسیله رابطه ذیل محاسبه کرد:

$$Entropy(t) = -\sum_j p(j|t) \log_2 p(j|t)$$

که در آن $p(j|t)$ میزان خلوص در دسته t از نود t می‌باشد. به این ترتیب الگوریتم بر روی مشخصه‌ای در هر مرحله انشعاب خواهد زد که بیشترین کاهش آنتروپی از نود مادر را ایجاد کرده و به عبارت بهتر بیشترین تمایز را از نود مادر در ایجاد خلوص دسته‌های حاصل ایجاد کند [۴۸]. شکل ذیل نمایی از درخت تصمیم ساخته شده بر روی خوشه مشتریان افراد تحصیلکرده را نشان می‌دهد.



شکل ۳ درخت تصمیم ساخته شده بر روی خوشه مشتریان افراد تحصیلکرده

به‌کارگیری درخت تصمیم در مقایسه با سایر تکنیک‌های پیش‌بینی کننده دارای این مزیت است که با کمک قواعد تصمیم‌گیری^{۱۴} می‌توان قواعد منتهی به ایجاد مقداری خاص برای متغیر هدف را پیگیری کرد. قواعد تصمیم‌گیری از روی درخت تصمیم به سادگی و با دنبال کردن یک مسیر از گره ریشه تا برگ‌ها ساخته می‌شوند. مجموعه این قواعد تصمیم‌گیری خود درخت را خواهند ساخت. قواعد تصمیم‌گیری به شکل اگر...، آن‌گاه... ساخته می‌شوند. بخش مقدم این رابطه شرطی شامل مقادیر مشخصه‌های مختلف در مسیر ساخته شدن آن قاعده و بخش تالی مربوط به مقدار متغیر هدف در گره برگ آن مسیر است. جدول ۶ بیانگر قواعد دسته‌بندی ساخته شده روی هر خوشه می‌باشد. دو عدد داخل پرانتز در مقابل هر قاعده به ترتیب از سمت راست نرخ



پشتیبانی و اطمینان آن قاعده را درخوشه مورد نظر بیان می‌کنند (نرخ پشتیبانی درصد رکوردهایی از هر درخت است که در گره برگ مسیر یک قاعده وجود دارند و نرخ اطمینان نیز درصد رکوردهایی از آن برگ را نشان می‌دهد که قاعده مورد نظر در آن صدق می‌کند). نتیجه ساخت این چهار درخت و قواعد تصمیم استخراجی از آن‌ها به شرح مندرج در جدول ۶ می‌باشد.

جدول ۶ قواعد رویگردانی استخراج شده برای مشتریان گروه‌های مختلف

		غیر رویگردان (۰/۶۷۵) (۰/۸۳۱)	بیش‌ترین فعالیت بانکی: بانک کشاورزی	گروه اول: مشتریان کاسب/ بازاری (نرخ دقت کلی: ۸۲/۹۲٪)
		رویگردان (۰/۷۵؛ ۰/۳۲۵)	بیش‌ترین فعالیت بانکی: بانک‌های دولتی دیگر- خصوصی	
		رویگردان (۰/۶۶۷؛ ۰/۰۵۸)	بیش‌ترین فعالیت بانکی: بانک‌های دولتی دیگر	
		رویگردان (۰/۵۰۰؛ ۰/۰۵۸)	بیش‌ترین فعالیت بانکی: بانک‌های خصوصی	گروه دوم: (افراد تحصیل‌کرده) (نرخ دقت کلی: ۸۲/۱۴٪)
	رویگردان (۰/۲۸۸) (۰/۷۳۳)	رضایت از طرز پاسخ‌گویی ≥ 8	بیش‌ترین فعالیت بانکی: بانک کشاورزی	
	غیر رویگردان (۰/۵۳۸) (۰/۸۲۱)	رضایت از طرز پاسخ‌گویی < 8		
		رویگردان (۱/۰؛ ۰/۱)	بیش‌ترین فعالیت بانکی: بانک‌های دولتی دیگر- خصوصی	گروه سوم: (کارمندان) (نرخ دقت کلی: ۹۰٪)
	غیر رویگردان (۰/۷۲) (۰/۹۱۷)	بزرگ‌ترین مشکل: مهارت کارکنان	بیش‌ترین فعالیت بانکی: بانک کشاورزی	
	رویگردان (۰/۷۷۸؛ ۰/۱۸)	بزرگ‌ترین مشکل: وام، دسته چک و سرعت پایین		
		رویگردان (۰/۷۵؛ ۰/۲)	بیش‌ترین فعالیت بانکی: بانک‌های دولتی دیگر- خصوصی	گروه چهارم: (زنان) (نرخ دقت کلی: ۹۰٪)
	رویگردان (۱/۰؛ ۰/۴)	رضایت از سود سپرده ≥ 6	بیش‌ترین فعالیت بانکی: بانک کشاورزی	
	رویگردان (۰/۸؛ ۰/۲۵)	رضایت از مکان بانک ≥ 8		
	غیر رویگردان (۱/۰؛ ۰/۱۵)	رضایت از مکان بانک < 8		

- خوشه اول (مشتریان کاسب و بازاری) برای این گروه از مشتریان، عمده‌ترین عامل تأثیرگذار بانکی است که در حال حاضر بیش‌ترین فعالیت بانکی خود را انجام می‌دهند و عموماً تمایلی به تغییر آن در آینده ندارند.
- خوشه دوم (افراد تحصیلکرده): بزرگ‌ترین گروه رویگردان در این خوشه مربوط به مشتریانی است که اگر چه بیش‌ترین فعالیت آن‌ها در حال حاضر در بانک کشاورزی صورت می‌گیرد ولی نارضایتی از طرز جواب‌گویی، کارکنان آن‌ها را به سمت سایر بانک‌ها سوق داده است.
- خوشه سوم (کارمندان): این گروه از مشتریان نیز چنان‌چه دارای حساب نزد سایر بانک‌های دولتی/خصوصی باشند، آن را به بانک کشاورزی ترجیح داده و بخشی از مشتریان که در حال حاضر دارای بیش‌ترین فعالیت نزد بانک کشاورزی هستند نیز در مواجهه با مشکلات مرتبط با اخذ دسته چک، وام و سرعت پایین عملیات به سمت سایر مؤسسات رویگردان خواهند شد.
- خوشه چهارم (زنان): در مورد این گروه نیز تعیین‌کننده‌ترین عامل رویگردانی دارا بودن بیش‌ترین فعالیت نزد بانک‌های دیگر بوده و در درجات بعدی نرخ سود سپرده‌ها و مکان بانک از عوامل رویگردانی ایشان هستند.

۶- اعتبارسنجی نتایج به دست آمده

اعتبارسنجی روش مدلسازی باید براساس داده‌های آزمون و نه داده‌هایی که در آموزش الگو به کار رفته است، بررسی شود. برای این کار روش‌های مختلفی توسعه داده شده است که از جمله آن‌ها می‌توان به تقسیم داده‌های اولیه به مجموعه‌های آموزش و آزمون با نسبت از پیش تعریف شده و اعتبارسنجی متقاطع n تایی اشاره کرد. در بسیاری از تحقیقات تلاش می‌شود تا در صورت امکان از داده‌های با تمایز بیش‌تر به عنوان مجموعه آزمون استفاده شود (به عنوان مثال داده‌های مربوط به چند ماه به عنوان آموزش و داده‌های بازه زمانی دیگری به عنوان آزمون به کار رود) تا از همبستگی احتمالی داده‌ها و تأثیر شرایط محیطی بر نتایج اجتناب شود.

در این تحقیق به دلیل محدودیت‌های موجود در نوع و نیز شیوه جمع‌آوری اطلاعات (به شکل پرسشنامه) استفاده از روش اعتبارسنجی متقاطع اجتناب شده و اقدام به جمع‌آوری



مجموعه‌ای مجزا از پرسشنامه‌ها در دو شعبه بانک به عنوان مجموعه آزمون شده است. از مجموع ۱۰۰ پرسشنامه توزیع شده در این دو شعبه، ۸۴ پرسشنامه تکمیل شده‌اند که مبنای اعتبارسنجی نتایج حاصل قرار گرفته‌اند. در شکل ذیل نتیجه ساخت مجدد الگو با این مجموعه داده و نتیجه کاربرد درخت C5.0 برای پیش‌بینی نرخ رویگردانی دیده می‌شود. نتیجه پیش‌بینی این الگو در ماتریس انطباق قابل مشاهده است.

Results for output field CHURNER		
Comparing \$C-CHURNER with CHURNER		
Correct	79	94.05%
Wrong	5	5.95%
Total	84	
Coincidence Matrix for \$C-CHURNER (rows show actuals)		
	F	T
F	53	2
T	3	26
Performance Evaluation		
F	0.386	
T	0.991	
Confidence Values Report for \$CC-CHURNER		
Range		0.415 - 1.0
Mean Correct		0.864
Mean Incorrect		0.672
Always Correct Above		1.0 (0% of cases)
Always Incorrect Below		0.416 (0.4% of cases)
95.18% Accuracy Above		0.0
2.0 Fold Correct Above		0.98 (68.85% of cases)

شکل ۴ اعتبارسنجی روش دسته‌بندی استفاده شده با تغییر ضرایب وزنی هزینه پیش‌بینی نادرست

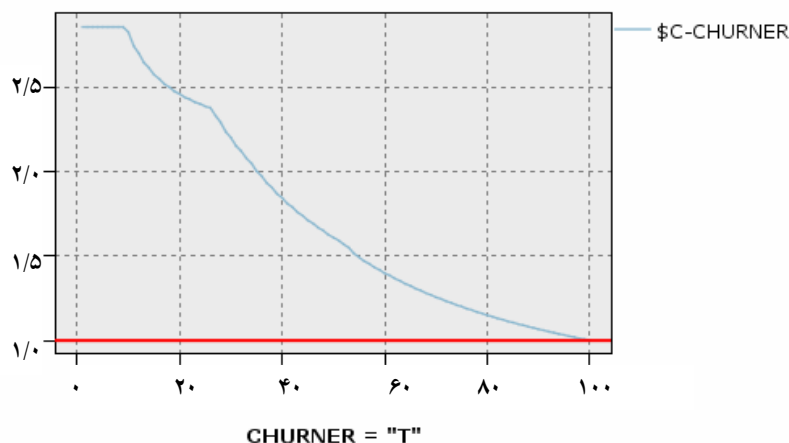
نرخ خطای کلی، حساسیت و تشخیص در این مورد به شکل زیر می‌باشد [۴۸]:

$$\text{نرخ خطای کلی} = \frac{(3+2)}{84} = 5.95\%$$

$$\text{sensitivity} = \frac{TP}{TP + FN} = \frac{26}{26 + 3} = 0.896$$

$$\text{specificity} = \frac{TN}{TN + FP} = \frac{53}{53 + 2} = 0.964$$

نمودار lift به‌ازای داده‌های آزمون در این الگو به شرح زیر است:



شکل ۵ نمودار lift به‌ازای اجرای الگو بر روی داده‌های آزمون

به‌طور معمول گفته می‌شود چنان‌چه نرخ lift به‌ازای ۱۰ درصد اول از ۲ بیش‌تر باشد، آن الگوی دسته‌بندی در جامعه مورد بررسی قابل قبول می‌باشد [۳۵، صص ۱۶۴-۱۷۴]. با این فرض، روش به‌کار رفته برای دسته‌بندی مشتریان دارای اعتبار کافی است.

۷- نتیجه‌گیری

در این مقاله تلاش شده است تا با بررسی پدیده رویگردانی مشتریان حوزه حساب‌های جاری اشخاص بانک کشاورزی، عمده‌ترین دلایل شکل‌گیری این پدیده مورد بررسی قرار گیرد. نخست در طول آن نرخ تخمینی رویگردانی هر مشتری با کمک مشخصه‌های مورد استفاده برای اندازه‌گیری آن به شکل برجسب رویگردانی مشتری در ادبیات تخمین زده شد سپس اقدام به خوشه‌بندی مشتریان مورد مطالعه به خوشه‌های مختلف دموگرافیک شد و در ادامه تأثیر مؤلفه‌های رضایت‌مندی و موانع تعویض تأمین‌کننده بر تصمیم رویگردانی مشتریان هر خوشه مورد ارزیابی قرار گرفت. از جمله نوآوری‌های این تحقیق می‌توان به ترکیب تحقیق پیمایشی با داده‌کاوی به منظور استخراج دلایل رویگردانی اشاره کرد که



تاکنون در مقاله مشابه داخلی به این موضوع پرداخته نشده است، نوآوری دیگر این تحقیق نیز به طور اساسی پرداختن به موضوع رویگردانی مشتری از خدمات بانکداری در ایران است که تاکنون تحقیق مشابهی در این زمینه صورت پذیرفته است.

نتایج حاصل از اجرای این تحقیق بیانگر تعیین‌کننده بودن مشخصه‌های دموگرافیک در عوامل تأثیرگذار بر وفاداری مشتریان رده‌های مختلف بوده است، به‌گونه‌ای که برای مشتریان از نوع کاسب/ بازاری تعیین‌کننده‌ترین عامل، بانکی است که در حال حاضر محل بیش‌ترین فعالیت بانکی ایشان می باشد (هزینه تعویض تأمین‌کننده، تعیین کننده‌ترین عامل برای ادامه روابط بوده است) و تأثیر مؤلفه‌های رضایت‌مندی بر تغییر نرخ وفاداری ایشان بسیار ناچیز بوده است. برای خوشه مشتریان تحصیلکرده، رضایت‌مندی از شیوه پاسخ‌گویی کارکنان برای خوشه کارمندان نحوه اخذ دسته چک و وام و برای خوشه زنان نرخ سود سپرده‌های پشتیبان و محل بانک تعیین‌کننده‌ترین عوامل وفاداری شناخته شده‌اند. با توجه به نتایجی که از این مقاله به دست آمده است، بانک کشاورزی برای حفظ مشتریان خود و اجرای برنامه‌های وفاداری آینده باید به بخش‌های مختلف بازار نگاهی متمایز داشته باشد. به این ترتیب که برای مشتریان رده کاسب/بازاری که بزرگ‌ترین و در عین حال مهم‌ترین گروه مشتریان حوزه حساب‌های جاری اشخاص هستند، اجرای برنامه‌هایی برای درگیرتر کردن آن‌ها در ارتباطات فعلی با بانک و بهره‌مندی این مشتریان با انواع خدمات ممکن به‌گونه‌ای که تعویض تأمین‌کننده را در آینده برای ایشان دشوارتر کند، بهترین رویکرد در مقابله با رویگردانی این گروه است. برای مشتریان رده تحصیلکرده طرز پاسخ‌گویی کارکنان تأثیرگذارترین عامل در بروز رویگردانی می‌باشد.

بنابراین ارتقای فرهنگ مشتری‌مداری و طرز برخورد کارکنان از طریق برنامه‌های آموزشی برای شعبی که دارای درصد زیادی از مشتریان رده تحصیلکرده هستند، راهکاری مناسب برای مقابله با رویگردانی این گروه است. برای رده مشتریان کارمند نیز بهبود روند اعطای دسته چک و تسهیل در شرایط اعطای وام هم جزء عوامل تأثیرگذار بر پیشگیری از بروز رویگردانی است که در صورت لزوم با سرمایه‌گذاری در این بخش‌ها می‌توان از رویگردانی این گروه جلوگیری کرد. روند رویگردانی گروه زنان نیز با افزایش تعداد شعب

بانک کشاورزی در مراکز مختلف شهر و نیز نرخ سپرده‌های پشتیبان در مقایسه با این نرخ در سایر بانک‌ها قابل کاهش می‌باشد.

در تحقیق حاضر، جامعه آماری تنها به ۱۰ شعبه بزرگ و تجاری اصلی بانک کشاورزی در شهر تهران محدود بوده است که در طول تحقیقات آینده می‌توان با توسعه جامعه آماری مزبور نتایج جامع‌تری کسب کرد. همچنین انواع تکنیک‌های بخش‌بندی و پیش‌بینی کننده نیز قابل استفاده در مراحل بخش‌بندی مشتریان و شناسایی روند رویگردانی ایشان هستند که می‌تواند بر روی داده‌های اولیه اجرا شده و نتایج آن‌ها با یکدیگر مقایسه شود. بهره‌گیری از داده‌های تراکنشی مشتریانی هم که به عنوان نمونه تصادفی انتخاب شده‌اند، می‌تواند سبب بررسی روند تراکنش‌های فوق و استفاده از آن درخوشه‌بندی مشتریان به کسب نتایجی جامع‌تر درخصوص روندهای رو به افول روابط با مشتریان و بررسی همزمان نظرات ارائه شده به وسیله این دسته از مشتریان شود و نتایجی بسیار مهم را در درک روابط منجر به رویگردانی برای بانک کشاورزی ایجاد کند که اجرای آن به وسیله پژوهش‌های آینده پیشنهاد می‌شود.

۸- پی‌نوشت‌ها

1. Voluntary churn
2. Non- voluntary churn
3. پسورد
4. Identification of the best data
5. Data Semantics
6. Feature selection
7. Development of predictive model
8. Validation of results
9. Multi dimensional scaling
10. Overall accuracy
11. Profit
12. Information gain
13. Entropy reduction
14. Decision rule



۹- منابع

- [1] Anderson J. L., Jolly L. D. Fairhurst A. E. ;"Customer relationship management in retailing: A content analysis of retail trade journals"; *Journal of Retailing and Consumer Services* 14:6: pp. 394-399, 2007.
- [2] Ang L. , Buttle F.;"Customer retention management processes: A quantitative study";*European Journal of Marketing*, 40:1-2: pp. 83-99, 2006.
- [3] Boone D. S. , Roehm M.; "Retail segmentation using artificial neural networks", *International Journal of Research in Marketing*, 19:3: pp. 287-301, 2002..
- [4] Hsieh N.-C.;"An integrated data mining and behavioral scoring model for analyzing bank customers"; *Expert Systems with Applications*, 27:4: pp. 623-633, 2004.
- [5] Minami C. , Dawson J.;"The CRM process in retail and service sector firms in Japan: Loyalty development and financial return"; *Journal of Retailing and Consumer Services*, 15:5: pp. 375-385, 2008.
- [6] Ngai E. W. T., Xiu, L. Chau, D. C. K. ;"Application of data mining techniques in customer relationship management: A literature review and classification"; *Expert Systems with Applications*, pp. 2592-2602, 2009.
- [7] Rygielski C., Wang J. C.,Yen D. C.; "Data mining techniques for customer relationship management"; *Technology in Society*, 24: pp. 483-502, 2002.
- [8] Chiang D. A., Wang Y. F., Lee S. L.,Lin C. J.; "Goal-oriented sequential pattern for network banking churn analysis"; *Expert Systems with Applications*, 25: pp. 293-302, 2003.
- [9] Bin L., Peiji S., Juan L.; "Customer churn prediction based on the decision tree in personal handy phone system service"; *First International Symposium on Data, Privacy and E-Commerce, China*, 8-11 July: pp. 8-11, 2007.

- [10] Buckinx W., Verstraeten G., Van den Poel D. ;"Predicting customer loyalty using the internal transactional database."; *Expert Systems With Applications* 32:1: pp. 125-134, 2007.
- [11] Keaveney S. M., Parthasarathy M.; "Customer switching behavior in online services: An exploratory study of the role of selected attitudinal, behavioral, and demographic factors"; *Journal of the Academy of Marketing Science*, 29:4: pp. 374-390, 2001.
- [12] Kim M. P. M., Jeong D.;"The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunications services"; *Telecommunications Policy*, 28: pp. 145-159, 2004.
- [13] Ling R. , Yen D. C. "Customer relationship management: an analysis framework and implementation strategies"; *Journal of Computer Information Systems*, 41: 3: pp. 82-97, 2001.
- [14] Qi J., Zhang Ya., Zhang Yi., Shi S.; "TreeLogit model for customer churn prediction"; IEEE Asia-Pacific Conference on Services Computing, Beijing, China, 2006.
- [15] Ng K. , Liu H.;"Customer retention via data mining"; *Artificial Intelligence Review*, 14:6: pp. 569-590, 2000.
- [16] Wang Y. F., Chiang D. A., Hsu M.-H., Lin C.-J., Lin I.-L. "A recommender system to avoid customer churn: A case study"; *Expert Systems with Applications*, 36: 4: pp. 8071-8075, 2009.
- [17] Wei C.P. Chiu I.;"Turning telecommunications call details to churn prediction: A data mining approach"; *Expert systems with applications*, 23: pp. 103-112, 2002 .
- [18] Yan L., Wolniewicz R. H., Dodier R. ;Predicting customer behavior in telecommunications; *IEEE Intelligent Systems*, 19:2: pp. 50-58, 2004.

- [19] Bloemer J. M. M., Brijs T., Vanhoof K., Swinnen G.; "Comparing complete and partial classification for identifying customers at risk" *International Journal of Research in Marketing*, 20:2: pp. 117-131, 2003.
- [20] Liao S.H., Chen J.-L., Hsum T.-Y.; "Ontology-based data mining approach implemented for sport marketing"; *Expert Systems with Applications*, 36:8: pp. 11045-11056, 2009.
- [21] Cymbalam A., Owczarczuk M., "Variables importance in questionnaire data on advertising" *Expert Systems with Applications*, 38:11: pp. 14218-14224, 2011.
- [22] Bae J. K., Kim J.; "Product development with data mining techniques: A case on design of digital camera"; *Expert Systems with Applications*, 38:8 :pp. 9274-9280, 2011.
- [۲۳] امیر شاهی ا.، سیاه تیری و.، روان بد ف.؛ "شناسایی عوامل مؤثر بر ایجاد "اعتماد" در مشتریان کلیدی بانک: مشتریان کلیدی بانک کارآفرین در شهر تهران"; *فصلنامه پژوهشهای مدیریت در ایران - مدرس علوم انسانی* ، ۱۳: ۴، صص ۶۱-۷۶ زمستان ۱۳۸۸.
- [۲۴] سهرابی ب.، خانلری ا.، آجرلو ن.؛ "الگویی برای تعیین چرخه عمر مشتریان (CLV) در صنعت بانکداری"; *فصلنامه پژوهشهای مدیریت در ایران - مدرس علوم انسانی* ، ۱۵: ۱، صص ۲۲۳-۲۳۹، بهار ۱۳۹۰.
- [25] Hadden J., Tiwari A., Roy R., Ruta D.; "Computer assisted customer churn management: State-of-the-art and future trends"; *Computers & Operations Research*, 34:10: pp. 2902-2917, 2005.
- [26] Kim H.-S., Yoon C.-H.; "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market" *Telecommunications Policy*, 28:9-10: pp. 751-765, 2004.
- [27] Oliver R. L.; "A cognitive model of the antecedents and consequences of satisfaction decisions"; *Journal of Marketing Research*, 17: 4: pp. 460-469, 1980.

- [28] Gupta S., Zeithaml V.; "Customer satisfaction and their impact on financial performance"; *Marketing Science*, 25: pp. 718-739, 2006.
- [29] Cronin J. Jr., Brady M., Hult G.; "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments"; *Journal of Retailing*, 76: pp. 93-218, 2000.
- [30] Rust R.T., Moorman C., Dickson P.R.; "Getting return on quality: Revenue expansion, cost reduction, or both?"; *Journal of Marketing*, 66 :4: pp. 7-24, 2002.
- [31] Gruca T., Rego L.; "Customer satisfaction, cash flow and shareholder value"; *Journal of Marketing* 69 :3: pp. 115-130, 2005.
- [32] Reichheld, F.F. Sasser Jr. W.E.; "Zero defections: Quality comes to services"; *Harvard Business Review* 68:5,: pp. 105-111, 1990.
- [33] Warrington P. T., Gangstad E., Feinberg R.; Multi-channel retailing and customer satisfaction: Implications for E-CRM, *marketing science*, 30: pp. 47-57, 2009.
- [34] Datta P., Masand B., ManD. R., Li B. ;"Automated cellular modeling and prediction on a large scale"; *Artificial Intelligence Review* 14:6: pp. 485-502, 2001.
- [35] Coussement K., Van den Poel D. ;"Integrating the voice of customers through call center emails into a decision support system for churn prediction"; *Information and Management*, 45: pp. 164-174, 2008.
- [36] Gladly N., Baesens B., Croux C.; "Modeling churn using customer lifetime value"; *European Journal of Operational Research*; 197: 1: pp. 402-411, 2009.
- [37] Verbeke W., Martens D., Mues C., Baesens B. ;"Building comprehensible customer churn prediction models with advanced rule induction techniques"; *Expert Systems with Applications*, 38:3: pp. 2354-2364, 2011.
- [38] Bock K. W. D., Van den Poel D.; "An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction"; *Expert Systems with Applications*, 38:10: pp. 12293-12301, 2011.

- [39] Shin D. H., Kim W. Y.; "Forecasting customer switching intention in mobile service: An exploratory study of predictive factors in mobile number portability"; *Technological Forecasting and Social Change*, 75: pp. 854-874, 2008.
- [40] Cooper D. R., Shindler P. S.; 'Business Research Methods' chapter 11: Survey methods; McGraw-Hill, Irwin, pp. 312-320, 2000.
- [۴۱] سرمد ز، بازرگان ع، حجازی ا؛ روش‌های تحقیق در علوم رفتاری (چاپ پانزدهم)، تهران: مؤسسه انتشارات آگاه، ۱۳۸۷.
- [42] Spiller J., Vlastic A., Yetton P.; "Post-adoption behavior of users of Internet Service Providers"; *Information and Management*, 44: pp. 513-523, 2007.
- [43] Liou J. J. H.; "A novel decision rules approach for customer relationship management of the airline market" ; *Expert Systems with Applications*, 36:3: pp. 4374-4381, 2009.
- [44] Burez J., Van den Poel D. ; "Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department"; *Expert Systems with Applications*, 35: pp. 497-514, 2008.
- [45] Cho Y. B., Cho Y. H., Kim S. H.; "Mining changes in customer buying behavior for collaborative recommendations, *Expert Systems with Applications*, 28: pp. 359-369, 2005.
- [46] Xie Y., Li X., Ngai E.W.T., Ying W.; "Customer churn prediction using improved balanced random forests"; *Expert Systems with Applications*, 36:3: pp. 5445-5449, 2009.
- [۴۷] رسول اف ج، سیفی م، رشیدی د؛ "شناسایی سطوح نیازها و عوامل مؤثر در رضایت مشتریان: پیش‌نیازی برای اصلاحات ساختاری در نظام بانکی"؛ *فصلنامه بانک کشاورزی*، ۱، صص ۱-۳۰، ۱۳۸۰.
- [48] Tan P-N., Steinbach M., Kumar V., Introduction to data mining, United States, Addison- Wesley Pub, 2005.