

بررسی ویژگی‌های رفتاری و کارکردی مشتریان شرکت مخابرات با رویکرد مدیریت ارتباط با مشتری با استفاده از روش‌های مبتنی بر داده کاوی

فاطمه احمدی آبکناری^{۱*}، شیرین احدزاده قناد^۲

۱- استادیار گروه فناوری اطلاعات و مهندسی کامپیوتر، دانشگاه پیام نور رشت

۲- کارشناس ارشد- مهندسی کامپیوتر، نرم افزار، دانشگاه آزاد رشت

چکیده

با رقابتی شدن صنعت مخابرات و رشد انتظارات مشتریان همگام با پیشرفت فناوری‌های ارتباطی، الزام تعیین و تبیین استراتژی‌های جدید ارتباط با مشتری در این صنعت احساس می‌شود. از راهکارهای وصول به این منظور، دسترسی به اطلاعات جامع از مشتریان برای شناخت مؤثر آنها است تا بتوان خدمات متناسب با ویژگی‌های کارکردی و رفتاری هر دسته را برای بهبود مدیریت ارتباط با مشتری ارائه داد. در این مقاله از اطلاعات پرداخت و بدهکاری مشترکین اداره مخابرات استان گیلان، شهرستان رشت در یک دوره شش ماهه استفاده شده است. اطلاعات مشتریان به صورت خاص خوشه‌بندی و تحلیل *RFM* شده و پس از شناخت خوشه‌های مختلف مشتریان و ارزیابی خوشه‌بندی بهینه آنها، با استفاده از ماتریس سودآوری وفاداری، از میان خوشه‌بندی‌های صورت گرفته، دسته‌بندی مشتریان هر خوشه انجام شده است. در این پژوهش، داده‌های مشتریان بر اساس روش تحلیل *RFM* انتخاب و دسته‌بندی شده و سپس با استفاده از الگوریتم خوشه‌بندی *K-Means*، سه خوشه عمده از مشتریان شناسایی شده‌اند و بر اساس ویژگی‌های رفتاری و الگوی مصرف هر خوشه، بسته پیشنهادی خدمات به آنها ارائه شده است. این بسته به منظور بهینه‌سازی مدیریت ارتباط با مشتری و ارائه خدمات بهتر در جهت افزایش ارزش حیات مشتری ارائه شد.

واژه‌های کلیدی: مدیریت ارتباط با مشتری، تحلیل *RFM*، ارزش حیات مشتری، خوشه‌بندی مشترکین.

۱-۱- مقدمه

ظهور فناوری‌های جدید، گسترش رقابت تجاری و دگرگونی‌های جدید اجتماعی و اقتصادی عصر حاضر، تغییرات چشمگیری را در فضای بازارها ایجاد کرده است. افزایش قدرت انتخاب مشتریان و عدم موازنه عرضه و تقاضا از جمله این تغییرات است.

این شرایط تغییر رویکرد سازمان‌ها را از ی انبوه^۱ به ی تک به تک^۲ به همراه داشته و با اتکا به سفارشی سازی خدمات و محصولات، انعطاف‌پذیری در ارائه کالاها و خلاقیت در تبلیغات ی، فعالیت می‌کنند.

تحت مفاهیم ی رابطه‌ای^۳، مشتری محوری و به تبع آن مدیریت ارتباط با مشتری و فرایندهای ارزش آفرینی برای مشتری از اهمیت ویژه‌ای برخوردار است (مؤتمنی و همکاران، ۱۳۹۰). مدیریت ارتباط با مشتری، الگویی جامع و یکپارچه است که با تغییر فرهنگ سازمانی و تحت پوشش قرار دادن تمامی فرایندهای سازمان، تغییر نگرش سازمان‌ها را از محصول محوری به مشتری محوری موجب می‌شود و مشتری دارایی ارزشمند سازمان‌ها محسوب می‌شود. از این رو سازمان‌های تجاری از یکسو برای تطابق با تغییرات و از سوی دیگر به جهت جذب مشتریان جدید و حفظ مشتریان موجود، بودجه‌های کلانی را برای مدیریت ارتباط با مشتری اختصاص می‌دهند (اشتیاقی و همکاران، ۱۳۹۴).

گردآوری اطلاعات مشتریان برای افزایش وفاداری از یک طرف، و حفظ آنها به منظور تکرار خرید برای سازمان اهمیت ویژه‌ای دارد؛ همچنین تمایز بین مشتریان، بررسی و تحلیل نیازهای آنها و ارائه خدمات یا محصولات متناسب با نیازها، از اهمیت کلیدی

برخوردار است. در این بین، شناخت الگوهای خرید و وفاداری مشتریان موجب شده تا سازمان‌ها خدمات و محصولات سفارشی ارائه دهند (صنایعی و همکاران، ۱۳۹۴).

سازمان‌های تجاری، روزانه حجمی وسیع از داده‌های تراکنش مشتریان را ذخیره می‌کنند. از آنجا که تحلیل این انباره‌های داده‌ای به صورت دستی بسیار دشوار است از پتانسیل ابزارهای فناوری اطلاعات، مانند تکنیک‌های داده کاوی و روش‌های پیش‌پردازش انباره داده استفاده می‌کنند تا داده‌ها را به دانشی سودمند و مورد نیاز برای سازمان تبدیل کنند. داشتن این اطلاعات پاسخگوی نیازهای دانشی سازمان‌ها است تا بتوانند رفتار مشتریان را بهتر شناسایی و پیش‌بینی کنند و خدمات رسانی مؤثرتر به گروه‌های مختلف مشتریان داشته باشند و در نهایت سودآوری را افزایش دهند (ابوعلی و همکاران، ۲۰۱۰)^۴ چرا که از چالش‌های اصلی سازمان‌های تجاری از دست دادن مشتریان با ارزش است. یک مشتری رویگردان که شرکت رقیب را برای ارائه خدمات برمی‌گزیند، هزینه‌های بیشتری را به شرکت تحمیل می‌کند؛ چرا که در مقایسه با بدست آوردن مشتری جدید، به علت عدم برگشت سرمایه، هزینه بیشتری را به شرکت تحمیل می‌کند. به منظور مقابله با این مشکل، درک رفتار مشتریان و طبقه‌بندی آنها امری سودمند است.

می‌توان با توجه به میزان استفاده مشتریان از سرویس‌های موجود و بهبود خدمات ارائه شده به هر طبقه، تأثیر شایانی در بهبود سودآوری مشتریان داشت. در این پژوهش تحلیل داده‌های مشترکین سازمان مخابرات برای مدیریت ارتباط مؤثر با مشتری با هدف خوشه‌بندی مشتریان بررسی می‌شود. برای شناسایی

1 Mass Marketing

2 One to One Marketing

3 Relationship Marketing

رفتار مشتری و تشخیص تقلب از حوزه‌های اصلی مدیریت ارتباط با مشتری (CRM) در صنعت مخابرات است (رحمانی و تقوی طلب، ۱۳۹۱)، (باراگوین و همکاران، ۲۰۰۱)^۱. تکنیک‌های رتبه‌بندی و دسته‌بندی مشتریان به مدیران امکان خواهد داد تا با هر مشتری بر اساس علایق و سلیقه‌های او برخورد کنند و از ویژگی‌های مربوط به هر دسته برای اتخاذ استراتژی‌های متناسب بهره‌گیرند. کاربرد داده کاوی در مدیریت ارتباط با مشتری در موارد زیر خلاصه می‌شود (مؤتمنی و همکاران، ۱۳۹۰) بدست آوردن مشتریان جدید، تقسیم‌بندی مشتریان، ی شخصی، تجزیه و تحلیل حفظ و ریزش مشتری و تجزیه و تحلیل و پیش‌بینی سودآوری مشتری.

پژوهش‌های زیادی در تحلیل داده‌ها برای خوشه‌بندی مشتریان انجام شده است. این تحلیل‌ها اعم از داده‌های جمعیت‌شناختی یا تراکنشی در سازمان‌های تجاری هستند که با استفاده از الگوریتم‌های داده کاوی بررسی شده‌اند (زین‌العابدینی و همکاران، ۱۳۹۲)، (کفاش پور و همکاران، ۱۳۹۱). در این بخش، به چند پژوهش صورت گرفته از مشتریان شرکت‌های مخابرات هم با استفاده از داده‌های تراکنشی آنان و هم با استفاده از داده‌های جمعیت‌شناختی پرداخته می‌شود. در پژوهش‌های مورد اشاره، بر تحلیل داده‌های مشتریان شرکت‌های مخابرات با استفاده از روش تحلیل RFM و همچنین دسته‌بندی گروه‌های مشتریان و یا خوشه‌بندی آنان در قالب خوشه‌های رفتاری گوناگون با استفاده از تکنیک‌های داده کاوی تمرکز شده است.

ایسونند و حسینی (۱۳۹۳) در پژوهشی برای بررسی وضعیت ارزش حیات مشترکین تلفن ثابت ایران به جمع‌آوری، آماده‌سازی و خوشه‌بندی داده‌ها بر اساس

رفتار مشترک، مشتریان در هر خوشه دسته‌بندی می‌شوند و بسته‌های پیشنهادی برای خدمت رسانی بهتر به هر خوشه با توجه به خصایص رفتاری آنان ارائه می‌شود.

به این منظور از داده‌های مشترکین شرکت مخابرات استان گیلان، شهر رشت در سه دوره دو ماهه و مجموعاً یک بازه شش ماهه استفاده شده است. و داده کاوی بر روی داده‌ها، با نرم افزار RapidMiner انجام شده که جزئیات الگوریتم‌ها و نتایج پژوهش در بخش ۳ ارائه می‌شود.

ساختار مقاله شامل بخش‌های زیر است: پس از مقدمه، در بخش دوم، به بررسی مطالعات پیشین در حوزه داده کاوی بر روی اطلاعات مشترکین مخابرات می‌پردازد. در بخش سوم، روش پژوهش بیان می‌شود. در بخش چهارم، با توجه به نتایج حاصل از خوشه‌بندی و بخش‌بندی مشترکین بسته‌های پیشنهادی هر خوشه ارائه می‌شود.

۲-۱ ادبیات و پیشینه نظری پژوهش

شرکت‌های مخابراتی در طول زمان حجم زیادی از داده‌های مشتریان را جمع‌آوری کرده‌اند که در صورت کاوش صحیح، می‌توان از میان آنها، اطلاعات ارزشمند و الگوهای جذابی را برای مدیریت ارتباط با مشتری استخراج کرد. داده کاوی ابزار مفیدی برای این کار است که دانش را از داده‌های حجیم با روش‌های خودکار و نیمه خودکار بدست آورده و می‌تواند پاسخگوی بسیاری از چالش‌ها در کسب و کار مخابراتی باشد. دسته‌بندی مشتریان بر اساس الگوهای رفتاری، شناسایی ویژگی‌های مشترک در خوشه‌های مشتریان، اتخاذ استراتژی‌های متناسب با هر خوشه، جلوگیری از کلاهبرداری، ارزیابی ارزش حیات مشتری، پیش‌بینی

عدم قطع خدمات، حتی در صورت وجود بدهی برای این مشترکین لحاظ شد. همچنین بسته‌های تشویقی متناسب با نیازهای هر گروه برای افزایش میزان رضایت مشترکین پیشنهاد شد.

نن و همکاران^۲ (۲۰۱۰) پژوهشی را با استفاده از داده کاوی برای مدلسازی مدیریت ارتباط با مشتریان شرکت مخابرات چین انجام دادند و به بخش‌بندی مشتریان براساس تقاضای آنها پرداختند. داده‌های مشتریان در مدل شامل عادات، نیازها، جنسیت و میزان درآمد می‌باشد.

در این پژوهش از الگوریتم خوشه بندی *K-Means* مبتنی بر مرکز استفاده شده است که تقسیم‌بندی کیفی مشتریان بر اساس ارزش حیات مشتری انجام می‌شود. بر این اساس مشترکین به سه گروه کاربران کلیدی، مهم، و کاربران عمومی تقسیم شدند. سیاست شرکت مخابرات در برخورد با این مشتریان، به ترتیب، حفظ مشتری و جلب رضایتمندی، درک نیازهای مشتریان به منظور برقراری ارتباط مستحکم‌تر با آنها و در نهایت نیازسنجی بوده است.

لی و همکاران^۳ (۲۰۱۱) در چین به بررسی استراتژی ترکیبی برای نگهداری مشتری در یک شرکت مخابراتی پرداختند. بر این پژوهش، داده‌های شش ماهه مشتریان در قالب داده‌های آموزشی انتخاب شد و در ارزیابی مدل از روش *Cross Validation* و خوشه بندی *K-Means* از نرم افزار *SPSS* استفاده شد. انتخاب مقدار *K* در تعیین استراتژی رویگردانی، حائز اهمیت است، چراکه در صورت کوچک بودن *K*، شناسایی رویگردانی مشکل خواهد بود و در صورت بزرگ بودن مقدار *K*، سازمان مجبور به طراحی استراتژی

تحلیل *RFM* پرداختند. تحلیل *RFM* از روش‌های مطرح در تحلیل رفتار مشتریان است که بر اساس سه عامل تازگی تراکنش، تکرار خرید، ارزش و حجم تراکنش عمل می‌کند این روش تنها بر حجم خرید تمرکز ندارد، بلکه تکرار تراکنش‌های مالی و فاصله زمانی آن با زمان تحلیل را در نظر می‌گیرد. خوشه‌بندی در این پژوهش بر اساس رفتار کارکردی انجام شده است که شامل شاخص ارزش کارکرد و میانگین مبلغ کارکردی هر مشترک در دوره مورد بررسی برای سازمان است. شاخص تکرار کارکرد براساس تعداد دوره کارکرد مشترک برای سازمان تعیین شد و شاخص تازگی کارکرد بر اساس آخرین دوره‌ای که مشترک کارکرد داشته است در نظر گرفته شد.

رفتار بدهکاری شاخص ضدا ارزش بدهی است که به صورت میانگین بدهی مشترک در دوره‌ها و آخرین دوره‌ای که مشترک بدهی داشته است، تعیین شد و شاخص تازگی بدهی که بر اساس آخرین دوره‌ای که مشترک بدهکاری داشته است، تعیین شد. در این پژوهش، گروه‌های رفتاری مشتریان در شش خوشه کارکردی و سه خوشه بدهکاری مشخص شد. به منظور سیاست‌گذاری برای حفظ بیشتر مشتریان و ارائه راهبرد بهتر، بر روی میزان متوسط درآمد ایجاد شده توسط یک مشتری (*ARPU*)^۱ تمرکز شد که در شرکت‌های مخابراتی داخلی عموماً هر گروه درآمدی به دو دسته *ARPU* مستقیم (درآمد مستقیم سازمان) و *ARPU* غیرمستقیم (در قالب سرویس‌های ارزش افزوده تقسیم‌بندی شدند).

از جمله می‌توان به *ADSL*، اینترنت آسان و اینترنت هوشمند اشاره کرد. با توجه به اینکه گروه‌های پرکارکردتر برای سازمان سودآورترند، تسهیلاتی چون

2 Nan and Shao

3 Li, X., Huang, Li, S. and Zhang

1 ARPU: Average Revenue Per User/Unit

نگهداری متفاوتی برای تعداد زیادی مشتریان خواهد شد که سربار هزینه توسعه را به دنبال خواهد داشت. داده‌ها در چهار گروه براساس نوع بسته خدمات، تعداد شکایات، متوسط طول تماس، تعداد تماس و تعداد پیامک خوشه‌بندی شدند. استراتژی مشترک بین تمامی گروه‌های رویگردان، تعداد شکایات بود. شناسایی نیازهای این گروه از طریق تحلیل و بررسی شکایات در جلوگیری از رویگردانی و ریزش مشتری پیشنهاد شد. یافته‌های این پژوهش نشان می‌دهد، حفظ مشتری به نسبت سایر فاکتورهای رقابتی، بیشترین اثرگذاری را در سودآوری سازمان داشته است. چهارچوب طراحی و تکمیل استراتژی نگهداری مشتری، شامل پیش‌بینی رویگردانی مشتریان بوده است. تقسیم‌بندی مشتریان متمایل به رویگردانی بالا با استفاده از الگوریتم *K-Means* انجام شده است و بر مبنای خصوصیات مشتریان، تجزیه و تحلیل دلایل رویگردانی بر اساس نتایج خوشه‌ها و طراحی استراتژی نگهداری و برنامه‌های صورت گرفته است.

در پژوهش صوفیه (۱۳۸۸)، بر روی تشخیص تقلب در مشترکین مخابرات با استفاده از تکنیک‌های داده-کاوی تمرکز شد. انواع گوناگونی از تقلب در صنایع مخابراتی وجود دارد از جمله تقلب در آبونمان، استفاده از کارت‌های تلفنی غیر مجاز و... که علت بسیاری از این مشکلات، شکاف و ضعف فناوری در سیستم‌های مورد استفاده است. با بهره‌گیری از تکنیک‌های کشف دانش و داده‌کاوی امکان تشخیص تلاش‌های فریبکارانه با دقت بالا وجود دارد. در این پژوهش، به بررسی الگوی مصرف مشترک با ریزمکالمه و صورتحساب مشترک پرداخته شد و تحلیل آن با داده‌کاوی ماشین-بردار پشتیبان (*SVM*)، شبکه عصبی، درخت تصمیم و الگوریتم خوشه‌بندی *K-Means* صورت گرفت و پس از ارزیابی روش‌های مذکور، روش ماشین‌بردار پشتیبان با دقت ۸۸٪ به عنوان روش مطلوب‌تر برای تشخیص موارد تقلب انتخاب شد.

الهیلان و همکاران (۲۰۱۴)^۳، برای افزایش سوددهی و وفاداری مشتریان و جلوگیری از ریزش و رویگردانی

کرامتی و همکاران (۲۰۱۴)^۱، احتمال رویگردانی مشترکین تلفن‌های همراه را با استفاده از تکنیک‌های داده‌کاوی پیش‌بینی کردند. روش‌های بکار گرفته شده دسته‌بندی مشتریان در این تحقیق، الگوریتم‌های درخت تصمیم، شبکه عصبی، ماشین‌بردار پشتیبان و الگوریتم *K-Nearest Neighbors* بوده است.

در پژوهشی که منصور و همکاران (۲۰۱۰)^۲ بر روی تهدید تغییر اپراتور شبکه به دلیل فضای رقابتی حوزه خدمات تلفن همراه انجام دادند، بر ضعف پوشش آنتن شبکه همراه اول در بیشتر مناطق ایران تأکید شد. به منظور دسته‌بندی مشتریان از الگوریتم‌های

1 Keramati, A. Jafari-Marandi, R. Aliannejadi, M. Ahmadian, I. Mozzafari, M. Abbasi, U.
2 Mansouri, Saraee and Amirfattahi

۳-۱ روش پژوهش

هدف اصلی این پژوهش، بخش‌بندی و تحلیل رفتار مشتریان براساس داده‌های تراکنشی رفتاری (متغیرهای *RFM* کارکردی و *RFM* بدهکاری) است تا از این طریق خوشه‌های مشتریان برای کمک به تصمیم‌گیری در خصوص بهبود مدیریت ارتباط با آنها شناسایی شود و تحلیل رفتار مشتریان شرکت مخابرات استان گیلان مورد بررسی قرار گیرد. علاوه بر این، بخش‌بندی مشتریان بر اساس متغیرهای *RFM* رفتاری همچون کارکرد اخیر مشترک، فراوانی کارکرد و مبلغ کارکرد صورت گرفت و عواملی مانند بدهی اخیر مشترک، تکرار بدهی و مبلغ بدهی در نظر گرفته شد.

مجموعه داده‌ها مربوط به تراکنش‌های مشتریان شرکت مخابرات استان گیلان شهر رشت است که در بازه شش ماهه به صورت سه دوره دو ماهه از نیمه اول سال ۹۳ است.

انتخاب نمونه به صورت تصادفی صورت پذیرفته است تا نتایج بر مبنای داده‌های بدست آید. تعداد ۷۶۶۲۹ رکورد (مشترک) به صورت اولیه در نظر گرفته شد که پس از فرایند پیش‌پردازش، تعداد ۷۵۸۲۸ رکورد به عنوان ورودی نهایی مورد استفاده قرار گرفت. گام پیش‌پردازش شامل شناسایی و از بین بردن داده‌هایی دور افتاده^۸ به روش تک متغیره آماری (هیستوگرام) و حذف برخی رکوردها با مقادیر مفقود^۹، ادغام رکوردهای مربوط به یک مشتری و حذف برخی رکوردهای با مقادیر دارای اختلال^{۱۰} انجام شد. همچنین تغییر قالب ویژگی‌ها، همچون ویژگی کارکرد اخیر (*R*) و بدهی اخیر (*R_b*) از قالب تاریخ به قالب عددی به صورت در نظر گرفتن تعداد روز به

آنان به دسته بندی مشتریان شرکت مخابرات *PTX* در اندونزی پرداختند و برای شناخت بهتر گروه‌های مشترکین و ارائه خدمات و سرویس‌های متناسب با نیاز هر دسته پیشنهادهایی ارائه دادند. داده‌های پرداخت و بدهکاری مشترکین با *SPSS* تحلیل شد. خوشه‌های شناسایی شده مشترکین سودآور و مشترکین بدون سودآوری بوده‌اند. بر اساس این دسته‌بندی، بسته‌های پیشنهادی مختلفی، مانند افزایش آگاهی مشترکین از خدمات موجود، فروش متقابل، تبلیغات بیشتر و... ارائه شد.

لین^۱ (۲۰۰۷)، نن^۲ (۲۰۱۰)، یان هانگ^۳ (۲۰۱۱)، تیان^۴ (۲۰۱۱)، هوآنگو^۳ و همکاران (۲۰۱۱) و ییلو^۴ (۲۰۱۳) در پژوهشی به فواید استفاده از روش‌های خوشه‌بندی برای تفکیک مشتریان شرکت‌های مخابراتی پرداختند. آنها با توجه به ویژگی‌های رفتاری مشتریان توجه نشان دادند.

در پژوهشی دیگر، شوهای^۵ و همکاران (۲۰۱۲) استفاده از درخت تصمیم را به عنوان ابزاری در شناخت ویژگی‌های رفتاری مشتریان مورد توجه قرار دادند. لیو و همکاران (۲۰۰۶) و هومین^۶ و همکاران (۲۰۰۳) به بخش‌بندی مشتریان مخابرات پرداختند تا رفتار آنها را پیش‌بینی کنند. در پژوهش مینگ^۷ و همکاران (۲۰۰۴) علل رویگردانی مشتریان شرکت مخابرات مورد بررسی قرار گرفته و در پژوهش ووبین (۲۰۰۳) از روش *SWARM* برای بررسی ویژگی‌های رفتاری مشتریان مخابرات استفاده شده است.

1 Lin, Q.

2 Tian, Z.

3 Huan-guo, Z. and Wei, L.

4 Ye, L.

5 Shui Hua, H., Lu, Sh., and Leung, S.

6 Hui-min, M., Yin, H., and XIAO, W.

7 Ming, J., and Lin L.

8 Outlier Data

9 Missing Values

10 Noising Values

در جدول ۱ نشان داده شده است. بدیهی است از سایر فیلدها در دیتاست اولیه مانند نام مشترکین، آدرس و ... به علت عدم ارتباط با اهداف تحقیق صرف نظر شده است.

جای استفاده از تاریخ قبض و نرمال‌سازی مقادیر به روش *MinMax* و گسسته‌سازی ویژگی‌ها صورت گرفت. فیلدهای استفاده شده در این پژوهش از دیتاست اولیه به همراه شرح دوره مورد بررسی هر فیلد

۴-۱ تجزیه و تحلیل داده‌ها

جدول ۱- توصیف متغیرهای رفتاری RFM کارکردی و بدهکاری

شرح متغیر	نام متغیر	
R: دوره اخیر کارکرد مشترک در بازه دو ماهه	کارکرد اخیر مشترک	RFM کارکردی
F: فراوانی کارکرد در بازه شش ماهه	فراوانی کارکرد	
M: میانگین ارزش ریالی کارکرد مشترک در بازه شش ماهه	مبلغ کارکرد	
R _b : دوره اخیر بدهی مشترک در بازه دو ماهه	بدهی اخیر مشتری	RFM بدهکاری
F _b : فراوانی بدهکاری در بازه شش ماهه	تکرار بدهی	
M _b : میانگین ارزش ریالی بدهکاری مشترک در بازه شش ماهه	مبلغ بدهی	

مذکور، یکبار با داده‌های نرمال‌سازی شده و بار دیگر با داده‌های گسسته‌سازی شده انجام گرفت. در شکل-های ۱ و ۲ نمایی از اعمال عملگرهای خوشه‌بندی ذکر شده در نرم‌افزار *RapidMiner* بر روی داده‌های مورد بررسی نشان داده شده است. به منظور ارزیابی خوشه‌های بدست آمده توسط الگوریتم‌های خوشه‌بندی و یافتن خوشه‌بندی بهینه، از میان روش‌های ارزیابی خوشه‌بندی از دو شاخص مجموع مربع خطا^۱ (*SSE*) از رده *Item Distribution Performance* که بر مبنای نحوه توزیع کل داده‌ها در خوشه‌های ایجاد شده است و شاخص دیویس بولدین^۲ (*DB*) از رده *Clustering Distance Performance* که بر مبنای فاصله بیشینه‌ای بین خوشه‌های و فاصله کمینه درون خوشه‌ای است (صنعی آباده و همکاران، ۱۳۹۱)، (جیاوی و کمبر، ۲۰۰۶)^۳ استفاده شد. ارزیابی خوشه‌ها با متغیرهای

در مرحله مدل‌سازی داده‌ها، هدف ارائه مدلی است که با استفاده از آن بتوان به شناسایی و تحلیل رفتار مشترکین دست یافت و بخش‌بندی و گروه‌بندی مشترکین را انجام داد و در نهایت اتخاذ استراتژی‌های متناسب با هر گروه را عملیاتی کرد. در این قسمت، ابتدا به بخش‌بندی مشترکین با استفاده از تکنیک خوشه‌بندی پرداخته می‌شود تا هدف‌گذاری بر روی هر بخش (خوشه) شناخته شده از مشترکین از طریق مدل صورت گیرد. به منظور خوشه‌بندی مشترکین، از دو الگوریتم *K-Means* و *X-Means* تحت نرم افزار *RapidMiner* نسخه ۵٫۳ با فایل ورودی به فرمت *CSV* استفاده شد. الگوریتم خوشه‌بندی *K-Means* بر روی شش متغیر مورد بررسی در حوزه *RFM* کارکردی و *RFM* بدهکاری (*R*، *M*، *F*، *R_b*، *F_b*، *M_b*) با مقادیر مختلف *K* (۲، ۳، ۴، ۵، ۶، ۷، ۸ و ۹) اعمال شد. الگوریتم خوشه‌بندی *X-Means* با مقدار *K=4* بر روی شش متغیر فوق‌الذکر صورت گرفت. الگوریتم‌های

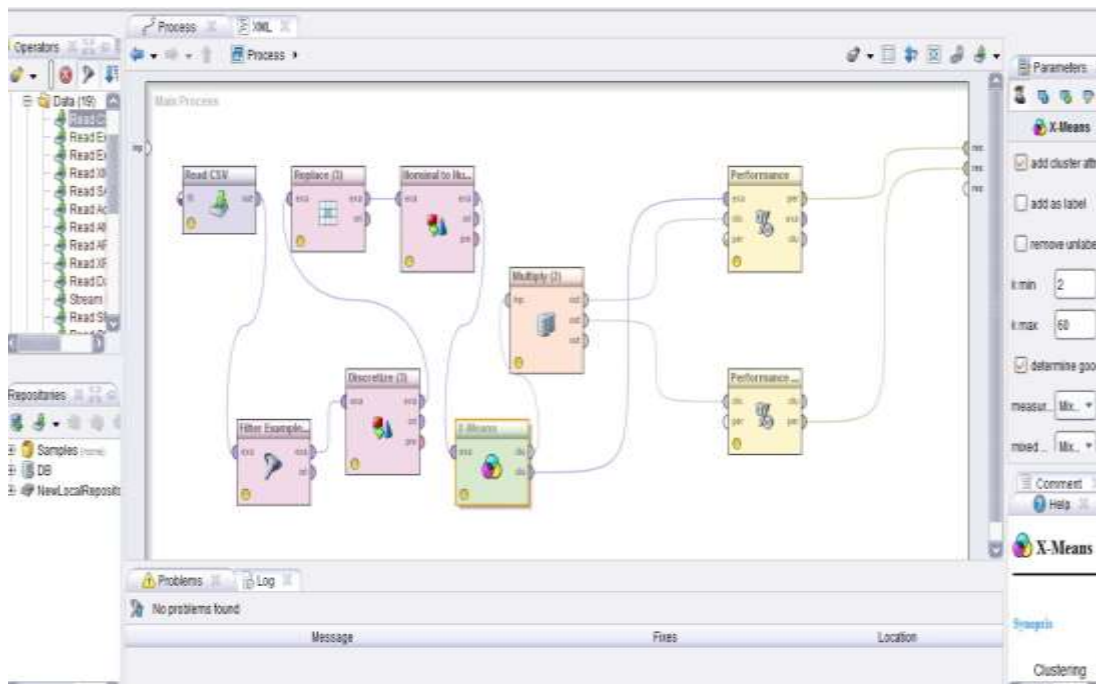
1 Sum of Squared Error

2 Davis Bouldin

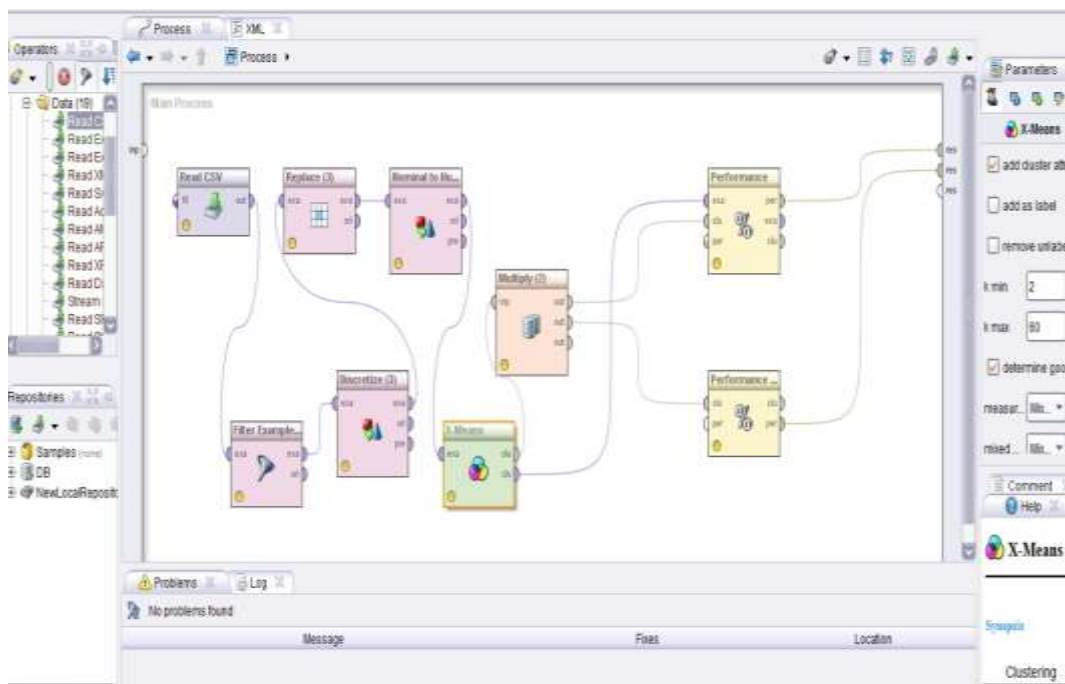
3 Jiawei and Kamber

ارزیابی‌های فوق در جدول ۲ و نمودارهای شکل‌های ۳ و ۴ و ۵ و ۶ نشان داده شده است. در شکل‌های ۳ و ۴ محور افقی مقدار K و محور عمودی مقدار شاخص DB و در شکل‌های ۵ و ۶، محور افقی مقدار K و محور عمودی مقدار شاخص SSE را نشان می‌دهد.

گسسته‌سازی شده و نرمال‌سازی شده با مقادیر مختلف ذکر شده K و بر روی دو روش خوشه‌بندی ذکر شده صورت پذیرفته و محاسبه شد تا روشی که کمترین مقدار SSE و DB را داشته باشد به عنوان بهترین خوشه‌بندی صورت گرفته مشخص گردد. نتیجه



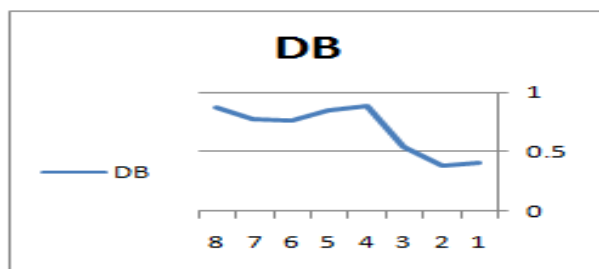
شکل ۱ - عملگر خوشه‌بندی K -Means مورد استفاده در نرم افزار RapidMiner



شکل ۲ - عملگر خوشه‌بندی X -Means مورد استفاده در نرم افزار RapidMiner

های خوشه‌بندی K -Means با $K=3$ برای داده‌های نرمال، X -Means با $K=4$ برای داده‌های نرمال، خوشه-بندی K -Means با $K=3$ برای داده‌های گسسته و X -Means با $K=4$ برای داده‌های گسسته هستند، به عنوان بهترین خوشه‌بندی‌های صورت گرفته شناخته می‌شوند.

با توجه به اینکه هرچه شاخص‌های ارزیابی مورد استفاده، مقادیر کمتری داشته باشند خوشه‌بندی با شرایط صورت گرفته دارای معیار جدایی خوشه‌ای مناسب تر و پیوستگی درون خوشه‌ای بهتری است، از این رو سطرهای پررنگ‌تر در جدول ۲ که الگوریتم-



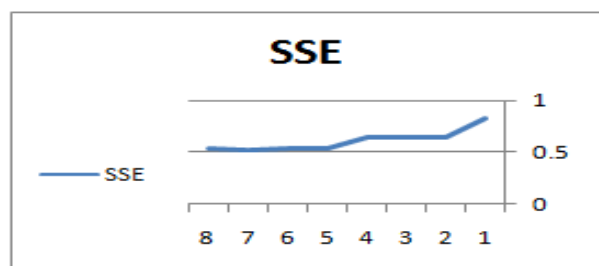
شکل ۳- نمودار مقادیر DB با داده‌های نرمال در روش K -Means



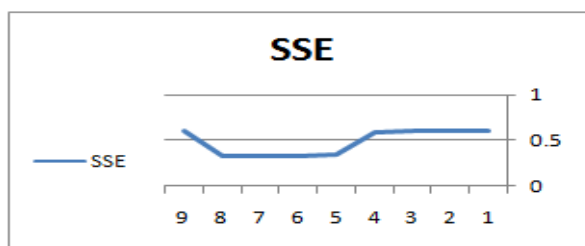
شکل ۴- نمودار مقادیر DB با داده‌های گسسته در روش K -Means

تعداد کل مشتریان را نشان می‌دهد. علت پوشش بالای خوشه صفر و فاصله مقدار آن نسبت به دو خوشه دیگر، جنبه انحصاری و غیر رقابتی شرکت مخابرات در بسیاری از خدمات است.

جدول ۳، توزیع مشترکین در خوشه‌های الگوریتم خوشه‌بندی K -Means با $K=3$ با داده‌های گسسته‌سازی شده (علت انتخاب داده‌های گسسته نسبت به داده‌های نرمال در ادامه شرح داده شده و در شکل‌های ۷ و ۸ نشان داده شده است) و درصد پوشش هر خوشه از



شکل ۵- نمودار مقادیر SSE با داده‌های نرمال در روش K -Means



شکل ۶- نمودار مقادیر SSE با داده‌های گسسته در روش K-Means

توجهی هستند، از این رو استفاده از داده‌های نرمال نمی‌تواند به وضوح تفاوت خوشه‌ها را نشان دهد، از این رو با توجه به شکل ۸ و تمایز بهتر داده‌های گسسته برای تفکیک خوشه‌ها از این داده‌ها استفاده می‌شود.

نمودارهای شکل‌های ۷ و ۸ تغییرات شش متغیر مورد بررسی در محور افقی را در هر یک از سه خوشه نشان می‌دهد. همانطور که در شکل ۷ مشخص است خوشه صفر و خوشه یک با یکدیگر همپوشانی زیاد داشته و خوشه صفر و دو نیز دارای همپوشانی‌های قابل

جدول ۲ - جدول مقایسه نتایج حاصل از ارزیابی الگوریتم‌های خوشه‌بندی

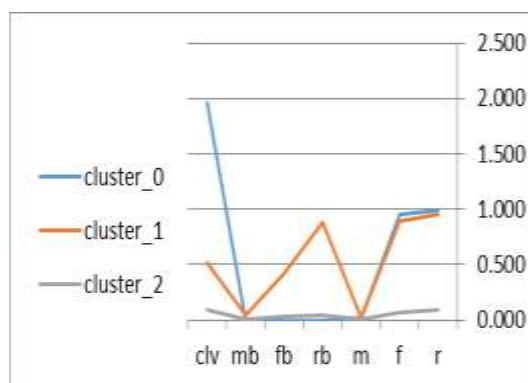
روش‌های پیش پردازش	الگوریتم خوشه‌بندی با مقادیر مختلف K	شاخص ارزیابی DB	شاخص ارزیابی SSE
داده‌های نرمال	K-Means - 2	0.41	0.83
	K-Means - 3	0.38	0.65
	K-Means - 4	0.54	0.65
	K-Means - 5	0.88	0.65
	K-Means - 6	0.84	0.53
	K-Means - 7	0.76	0.53
	K-Means - 8	0.77	0.52
	K-Means - 9	0.87	0.53
	X-Means - 4	0.5	0.65
داده‌های گسسته	K-Means - 2	1.2	0.61
	K-Means - 3	0.77	0.61
	K-Means - 4	1.01	0.61
	K-Means - 5	0.98	0.6
	K-Means - 6	0.94	0.34
	K-Means - 7	0.91	0.33
	K-Means - 8	0.9	0.33
	K-Means - 9	1.03	0.33
	X-Means - 4	1.01	0.61

جدول ۳- آمار تعداد و درصد مشترکین در هر خوشه (الگوریتم *K-Means* با $K=3$ با داده‌های گسسته‌سازی شده)

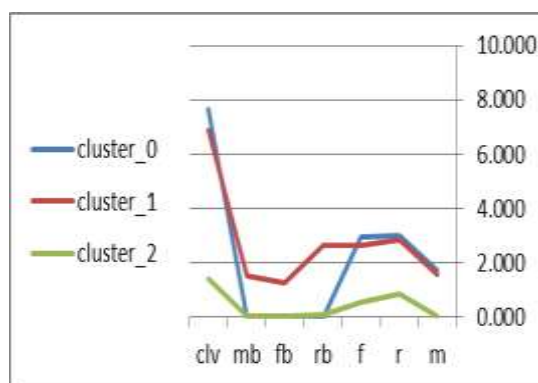
خوشه	تعداد مشترکین از خوشه	میزان مشترکین در هر خوشه (درصد)
صفر	۴۷۹۴۶	٪۶۳,۲
یک	۱۸۵۵۱	٪۲۴,۴
دو	۹۳۳۱	٪۱۲,۳
جمع	۷۵۸۲۸	۱۰۰

ارزش حیات مشتری (*CLV*) در جداول ۴ و ۵، از حاصل تفریق مجموع شاخص‌های بدهکاری از مجموع شاخص‌های کارکردی، طبق فرمول (۱) محاسبه شد:

$$CLV = M_{R'} + M_{F'} + M_{M'} - (M_{R'b} + M_{F'b} + M_{M'b}) \quad (1)$$



شکل ۷- متغیرهای RFM خوشه‌های کارکردی و بدهکاری با داده‌های نرمال‌سازی شده



شکل ۸- متغیرهای RFM خوشه‌های کارکردی و بدهکاری با داده‌های گسسته‌سازی شده

نتایج حاصل از رفتار کارکردی و بدهی در هر خوشه در جداول ۴ و ۵ به ارائه ویژگی‌های مشابه مشترکین هر

از آنجایی که اعضای قرار گرفته در یک خوشه، دارای ویژگی‌های رفتاری مشابهی هستند، با بررسی

خوشه و دسته‌بندی آنان با توجه به روش دسته‌بندی مشتریانی در قالب ماتریس سودآوری- وفاداری می‌پردازیم.

جدول ۴- متوسط ارزش شاخص‌های RFM در هر خوشه (با داده‌های گسسته‌سازی شده)

متغیرها	خوشه صفر	خوشه یک	خوشه دو
R	1. 731	1. 555	0. 015
F	2. 983	2. 830	0. 852
M	2. 925	2. 657	0. 550
R_b	0. 000	2. 650	0. 082
F_b	0. 000	1. 268	0. 046
M_b	0. 000	1. 523	0. 029
CLV	7. 638	1. 601	1. 261
نام دسته	مشتریان وفادار	مشتریان بدهکار (صدف)	مشتریان حساس (بیگانگان)

ماتریس سودآوری-وفاداری به عنوان یکی از روش‌های کلاسیک بخش‌بندی مشتریان می‌باشد که در قالب یکی از چهار گروه دوستان واقعی^۱، پروانه‌ها^۲، صدف‌ها^۳ و بیگانگان^۴ (موتمنی و همکاران، ۱۳۹۰) است.

جدول ۵- متوسط ارزش شاخص‌های RFM در هر خوشه (با داده‌های نرمال‌سازی شده)

متغیرها	خوشه صفر	خوشه یک	خوشه دو
R	۰,۹۸۸	۰,۹۵۰	۰,۰۹۴
F	۰,۹۵۶	۰,۸۹۲	۰,۰۷۰
M	۰,۰۲۳	۰,۰۲۲	۰,۰۰۲
R_b	۰,۰۰۰	۰,۸۸۲	۰,۰۴۴
F_b	۰,۰۰۰	۰,۴۱۸	۰,۰۳۰
M_b	۰,۰۰۰	۰,۰۴۷	۰,۰۰۲
CLV	۱,۹۶۷	۰,۵۱۷	۰,۰۹۰
نام دسته	مشتریان وفادار	مشتریان بدهکار (صدف)	مشتریان حساس (بیگانگان)

علت دسته‌بندی هر خوشه در قالب گروه‌های مشتریان در ماتریس سودآوری- وفاداری به شرح زیر است:

1 True friends
2 Butterflies
3 Barnacles
4 Strangers

جدول ۶- اقدامات پیشنهادی به منظور حفظ مشترکین هر خوشه

خوشه	اقدام پیشنهادی	اهداف (پیش‌بینی نتایج اقدام)
0	تشویق مشترکین خوش حساب و وفادار، بهبود ارتباط بیشتر با مشترکین از طریق نقاط تماس گوناگون (مرکز تماس پاسخگو، سایت، ایمیل و پیامک)، افزایش ساعات اینترنت رایگان در صورت استفاده از ADSL مخابرات، تسهیل در تخصیص امکانات ADSL در صورت تقاضا، بخشودگی پالس، شرکت در جوایز خوش حسابی مکالمات ارزان آخر هفته	افزایش رضایتمندی، افزایش سودآوری، بهبود کانال‌های ارتباطی
۱	ممانعت از قطع دو طرفه و دائمی، اعمال قطع یک طرفه آنها با فراهم نمودن امکان پرداخت اقساطی بدهی، تشویق مشترکین در صورت پرداخت اقساط در سر رسید مقرر	جلوگیری از رویگردانی مشترکین، سوق دادن مشترکین به سمت مشترکین سودآور، وصول مطالبات معوقه شرکت، افزایش درآمد و سودآوری شرکت
۲	ارائه چند دوره مکالمه رایگان، شرکت در قرعه‌کشی و سرویس‌های ارزش افزوده همانند اینترنت پرسرعت (با فرض ارائه این سرویس‌ها به منظور افزایش وابستگی بیشتر مشترک به سازمان)	سوق دادن مشترکین از ناپایدار به وفادار، جلوگیری از رویگردانی مشترکین، جلوگیری از ریزش مشترکین، افزایش درآمد و سودآوری شرکت

بر اساس ماتریس سودآوری-وفاداری در دسته صدف‌ها دسته‌بندی می‌شوند.

خوشه دو: مشترکین غیر وفادار و موقت محسوب می‌شوند و مقادیر F ، R و M و F_b ، R_b و M_b آنها پایین می‌باشد و نشانگر مشتریان کم کارکرد و یا بی کارکرد است (سازمان برای مشترکین بی کارکرد حداقل کارکرد ۲۰۰ پالس را منظور می‌نماید) و یا ممکن است مشترکین جدیدی باشند که اخیراً به جمع مشترکین پیوسته و بر اساس ماتریس سودآوری در دسته بیگانگان جای می‌گیرند. در اتخاذ سیاست‌های نگهداری از مشتریان این دسته، شرکت باید ترازش مناسبی بین

خوشه صفر: مشترکین وفادار و سودآور هستند، چراکه مقادیر F ، R و M آنها بالاتر از سایر گروه‌ها بوده و مقادیر F_b ، R_b و M_b آنها صفر می‌باشند. مشترکین دسته‌بندی شده در این خوشه، بر اساس ماتریس سودآوری مشتریان، در رده دوستان واقعی شرکت محسوب می‌شوند و به آنها مشترکین وفادار اطلاق می‌شود.

خوشه یک: مشترکین وفادار شرکت محسوب می‌شوند، اما سودآوری زیادی برای شرکت ندارند چراکه دارای مقادیر F ، R و M نسبتاً خوبی هستند، اما مقادیر F_b ، R_b و M_b نیز بالاتر از سایر خوشه‌ها می‌باشد.

فهرست منابع

۱- ایسوند، ف. حسینی، م. (۱۳۹۳). رویکردی نوین در بخش‌بندی مشتریان تلفن ثابت با بکارگیری مدل RFM. نشریه تخصصی مهندسی صنایع. ۴۸، ۳۶-۲۳، ۱.

۲- رحمانی، م. تقوی طلب، س. (۱۳۹۱). بهینه‌سازی مدیریت ارتباط با مشتری با استفاده از خوشه‌بندی فازی در شرکت‌های مخابرات. ششمین کنفرانس داده‌کاوی ایران. ۱۴۸. تهران. ایران.

۳- اشتیاقی، ف. نخعی کمال آبادی، ع. البدوی، الف. (۱۳۹۴). تخصیص بهینه بودجه‌های ارتباط با مشتری با رویکرد بهینه‌سازی استوار با هدف پیشینه‌سازی حقوق صاحبان سهام از ارزش مشتری (CE). فصلنامه علمی-پژوهشی تحقیقات ی نوین. سال پنجم، شماره اول، پیاپی (۱۶).

۴- صناعی، ع. حدادیان، ع. باقریه مهدی، الف. (۱۳۹۴). بررسی الگوی وفاداری مشتریان در فروشگاه‌های زنجیره‌ای ایران. فصلنامه علمی-پژوهشی تحقیقات ی نوین. سال پنجم، شماره اول، پیاپی (۱۶).

۵- صوفیه، ج. (۱۳۸۸). تشخیص فریبکاری در مخابرات با استفاده از داده‌کاوی. ماهنامه دنیای مخابرات و ارتباطات، ۶۴، ۴۴-۴۶.

۶- زین‌العابدینی، ف. مهدوی، م. خان بابایی، م. (۱۳۹۲). مدل بکارگیری تکنیک‌های داده‌کاوی در شناسایی، بخش‌بندی و تحلیل رفتار مشتریان خدمات بانکداری الکترونیک. فصلنامه علمی-پژوهشی تحقیقات ی نوین. ۳، ۲، ۱۷۵-۱۸۸.

۷- کفاش پور، آ. توکلی، ا. عزیزاده زوارم، ع. (۱۳۹۱). بخش‌بندی مشتریان بر اساس ارزش دوره عمر آنها با استفاده از داده‌کاوی بر مبنای مدل آر.

هزینه‌های نگهداری این دسته از مشتریان و عدم امکان برگشت سرمایه را در نظر داشته باشد. بسته پیشنهادی مناسب برای هر خوشه از مشترکین در جدول ۶ ارائه شده است:

۱-۵ نتیجه‌گیری

در این مقاله با بکارگیری الگوریتم‌های داده‌کاوی به منظور بهبود مدیریت ارتباط با مشتری، مشترکین شرکت مخابرات استان گیلان شهر رشت در سه خوشه تقسیم‌بندی شدند. و با توجه به بیشترین شباهت در الگوهای رفتاری گروه‌بندی شدند. سپس به بررسی ارزش عمر مشتری در هر گروه پرداخته شد و هر خوشه از نظر ماتریس سودآوری-وفاداری در گروه‌های مشتریان از دیدگاه این ماتریس قرار گرفت.

خوشه بندی مشترکین بر مبنای چارچوب ارائه شده در این تحقیق به شرکت مخابرات کمک می‌کند تا مشتریانی را که دارای الگوهای رفتاری مشابه برحسب متغیرهای کارکردی و بدهکاری هستند، در گروه‌های یکسان در نظر گرفته و سرویس متناسب با این ویژگی‌ها را به روش‌های متفاوت به مشتریان پیشنهاد و ارائه نماید و از ارائه روش کلاسیک و سنتی یک سرویس برای همه مشتریان که در دنیای امروز بی‌معناست بپرهیزد. البته تکمیل داده‌های جمعیت‌شناختی مشترکین شامل تحصیلات، وضعیت تأهل، شغل، جنسیت، سن، تاریخ تولد و... می‌تواند شرکت مخابرات را در ارائه سرویس مناسب به مشترکین بالقوه با توجه به نتایج اعمال الگوریتم‌های دسته‌بندی درخت تصمیم و الگوی شناسایی اقلام تکراری بر روی داده‌های مشترکین، یاری نماید.

- 17- Li, X., Huang, Y., Li, S. and Zhang, Y. (2011). Hybrid Retention Strategy Formulation in Telecom Based on K-Means Clustering Analysis. In Proceedings of the International Conference on E-Business and E-Government (ICEE). Shanghai, China.
- 18- Mansouri, R., Saraee, M. and Amirfattahi, R. (2010). Application of Data mining in Predicting Cell Phones Subscribers Behavior Employing the Contact Pattern. International Conference on Data Storage and Data Engineering (DSDE). Bangalore, India.
- 19- Nan, W. K. and Shao, Q. (2010). Application of Data Mining Technology in CRM. In Proceedings of the 2nd International Workshop on Database Technology and Applications (DBTA). Hubei, China.
- 20- Yun-Hung, G. (2011). Application of Improved K-Means Algorithm in Telecom Customer Segmentation. International Journal of Computer Simulation 8: 034.
- 21- Wu, B. (2003). A Customer Behavior Analysis Algorithm Based on SWARM Intelligence. Chinese Journal of Computers- Chinese Edition-26. 8: 913-918.
- 22- Ming, J., and Lin L. (2004) The Realization of Customer-Churn Model Based on DM. Computer Engineering and Applications, 4: 057.
- 23- Wang, Ch., Xiaoyun, H., and Biyan, W. (2003). An Empirical Study of the Relationship between Customer Satisfaction and Loyalty. Nankai Business Review, 4: 70-74.
- 24- Hui-min, M., Yin, H., and XIAO, W. (2003). Customer Potential Value Prediction and Customer Segmentation. Industrial Engineering and management, 2: 009.
- 25- Liu, Ying-zi, and Hao, WU. (2006). A Summarization of Customer Segmentation Methods. Journal of Industrial Engineering and Engineering Management, 1: 011.
- 26- Lin, Q. (2007). Mobile Customer Clustering Analysis based on Call Detail Records. Communications of the IIMA 7. 4: 95.
- اف. ام. پژوهش‌های مدیریت عمومی. پژوهش‌های مدیریت عمومی. ۵، ۱۵، ۶۳-۸۴.
- ۸- صنیعی آباده، م. محمودی، س. طاهرپرور، م. (۱۳۹۱). داده‌کاوی کاربردی. انتشارات نیاز دانش. چاپ اول.
- ۹- مؤمنی، م. فعال قیومی، ع. (۱۳۹۱). تحلیل آماری با استفاده از SPSS، انتشارات کتاب نو.
- ۱۰- مؤتمنی، ع، جعفری، الف و مجرد، ف. (۱۳۹۰). مدیریت ارتباط با مشتری. چاپ و نشر بازرگانی.
- ۱۱- احمدی آبکناری، ف. احدزاده قناد، ش. (۱۳۹۴). بهبود مدیریت ارتباط با مشتری در شرکت مخابرات با خوشه بندی و تحلیل آر. اف. ام مشترکین. دومین همایش ملی مهندسی کامپیوتر و فناوری اطلاعات، نجف آباد.
- 12- Keramati, A. Jafari-Marandi, R. Aliannejadi, M Ahmadian, I. Mozzafari, M. Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. International Journal of Applied Soft Computing, Elsevier. 10.1016/j.asoc.2014.08.041.
- 13- Alhilman, J. Moch, M, Wiyono, M. Margono, K. (2014). Predicting and Clustering Customer to Improve Customer Loyalty and Company Profit. In Proceedings if 2nd International Conference on Information and Communication Technology (ICOICT). Indonesia.
- 14- AbuAli, A. N. and Abu-Addose, H. Y. (2010). Data Warehouse Critical Success Factors. European Journal of Scientific Research, Vol. 42(2). pp: 326-335.
- 15- Baragoin, C., Andersen, Ch., Bayerl, S., Bent G., Lee J. and Schommer, Ch. (2001). Mining Your Own Business in Telecoms Using DB2 Intelligent Miner for Data. IBM Corporation. First Edition. California, USA.
- 16- Jiawei, H. and Kamber, M. (2006). Data Mining: Concepts and Techniques. Morgan Kaufmann Publication. Second Edition, Boston, USA.

- 29- Shui Hua, H., Lu, Sh., and Leung, S. (2102). Segmentation of Telecom Customers Based on Customer Value by Decision Tree Model. *Expert Systems with Applications* 39. 4: 3964-3973.
- 30- Ye, L. (2013). Customer Segmentation for Telecom with the K-Means Clustering Method. *Journal of Information Technology*: 12. 3: 409.
- 27- Tian, Z. (2011). Role Defining Using Behavior-Based Clustering in Telecommunication Network. *Expert Systems with Applications* 38.4: 3902-3908.
- 28- Huan-guo, Z. and Wei, L. (2011). Application of Telecom Customer Segmentation Base on Improved C-means Algorithm. *Computer Simulation* 6: 047.