



## توسعه مدل RFM به LRFM در خوشه بندی مشتریان مخابرات خراسان جنوبی

وحیده بابائیان

عضو هیات علمی دانشگاه صنعتی بیرجند،

[babaiyan@birjandut.ac.ir](mailto:babaiyan@birjandut.ac.ir)

سیده عطیه سرفرازی

کارشناس دانشگاه صنعتی بیرجند،

[sarfarazi.a.72@gmail.com](mailto:sarfarazi.a.72@gmail.com)

### چکیده

شرکت های مخابرات برای حفظ رابطه خوب با مشتریان موجود خود و جذب مشتریان جدید و شناسایی مشتریان سودده یا ضرر رسان، از تکنیک های داده کاوی استفاده می کنند. خوشه بندی موجب شناخت صحیح تر مشتریان می گردد تا در تعریف و تصمیم گیری طرح های تبلیغاتی، مورد استفاده قرار گیرد. در این مطالعه از سوابق خرید ۹۹۹ مشتری در شرکت مخابرات خراسان جنوبی که طی یکسال جمع آوری شده است استفاده کردیم. هدف از این پژوهش تقسیم بندی مشتریان به چند خوشه می باشد که پس از مشخص شدن خوشه ها و تعداد اعضای هر کدام، مشتریان پر مصرف را از سیستم خارج و مشتریان پر ارزش از دست رفته را نیز شناسایی کرد. در این تحقیق با نرم افزار کلمنتاین و توسعه مدل RFM به مدل LRFM و دو الگوریتم TwoStep و k\_Means توانستیم مشتریان را به پنج دسته ی وفادار، بلقوه، جدید، از دست رفته و پر مصرف تقسیم کنیم تا این بخش بندی معیار مناسبی در تصمیم گیری های آینده شرکت باشد و بتوانیم در خصوص هر دسته از مشتریان تصمیم گیری بهتری در آینده داشته باشیم.

واژگان کلیدی: مدل LRFM، الگوریتم TwoStep، الگوریتم k\_Means.

## مقدمه

داده‌کاوی به مفهوم استخراج اطلاعات نهان و یا الگوها و روابط مشخص در حجم زیادی از داده‌ها در یک یا چند بانک اطلاعاتی بزرگ است. بسیاری از مردم داده‌کاوی را مترادف واژه‌های رایج کشف دانش از داده‌ها می‌دانند. داده‌کاوی پایگاه‌ها و مجموعه‌های حجیم داده‌ها را در پی کشف و استخراج دانش، مورد تحلیل و کند و کاوهای ماشینی و نیمه‌ماشینی قرار می‌دهد. این گونه مطالعات و کاوش‌ها را در واقع می‌توان همان امتداد و استمرار دانش کهن و همه جا گیر آمار دانست. تفاوت عمده در مقیاس، وسعت و گوناگونی زمینه‌ها و کاربردها، و نیز ابعاد و اندازه‌های داده‌های امروزی است که شیوه‌های ماشینی مربوط به یادگیری، مدل‌سازی، و آموزش را طلب می‌نماید.

به طور کلی، داده‌کاوی فرآیند تجزیه و تحلیل داده‌ها از دیدگاه‌های مختلف و خلاصه آن را به اطلاعات مفید است. هدف از استخراج این اطلاعات مفید افزایش درآمد، کاهش هزینه‌ها، یا هر دو می‌تواند باشد. نرم افزارهای داده‌کاوی ابزارهای برای تجزیه و تحلیل داده‌ها هستند و اجازه می‌دهد تا کاربران به تجزیه و تحلیل داده‌ها از بسیاری از ابعاد و یا زوایای مختلف، بپردازند و روابط ناشناخته بین آن‌ها را کشف کنند. شناسایی مشتریان و نیازهای مشتریان از عوامل مهمی است که تولیدکنندگان برای به دست آوردن مزیت رقابتی در ارائه محصولات یا خدمات خود به مشتریان از آن کمک می‌گیرند. شرکت‌ها باید مشتریان خود را اولویت‌بندی نمایند و تمرکز خود را بر مشتریان کلیدی قرار دهند. اهمیت این موضوع به این دلیل است که با قطع ارتباط مشتریان و تمایل آن‌ها به رقبای شرکت نه تنها باعث ضرر و زیان مالی می‌گردد بلکه از دست دادن شهرت و اعتبار خوب را هم در پی دارد. همچنین ممکن است مشتریان فعلی تجارب منفی خود را با مشتریان بالقوه به اشتراک بگذارند و این منجر به از دست رفتن اعتماد مشتریان بالقوه می‌شود.

سازمان‌ها می‌توانند اهداف اساسی خود، از جمله دستیابی به مزیت رقابتی را با شناسایی گروه‌های مختلف مشتریان و نیازهای آن‌ها دنبال کنند تا با نیل به رضایت مشتری، که خود منجر به وفاداری مشتری در بلندمدت می‌گردد به سوددهی بیشتر برسند. همچنین شناسایی مشتریان کلیدی و حفظ آن‌ها نه تنها منجر به جذب مشتریان جدید می‌شود بلکه برای پر کردن جای خالی کسانی که تصمیم به قطع قرارداد با شرکت بسته اند نیز مفید است. چرا که که عمدتاً هزینه‌های کسب مشتری کلیدی جدید پنج برابر بیشتر از هزینه حفظ مشتری فعلی می‌باشد (کاتلر و آرمسترانگ، ۱۳۸۶).

طبقه‌بندی یکی از موضوعات مهم در مدیریت ارتباط با مشتری است. در طبقه‌بندی مشتریان، کل جمعیت مشتریان را به گروه‌های کوچکتری تقسیم می‌کنیم، به طوری که مشتریان یک گروه دارای ویژگی‌های مشابه هم هستند. در حالت ایده آل، سازمان باید یک درک خوب در مورد همه مشتریان خود داشته باشد، که این امر در دنیای واقعی امکان‌پذیر نیست. طبقه‌بندی مشتریان و خوشه‌بندی آنان به مدیران کمک می‌کند که درک بهتری از نیازهای آنان داشته باشند. داده‌کاوی دارای ابزارهای مختلف برای طبقه‌بندی مشتریان است که از تکنیک‌های متعددی بهره می‌گیرد. معروف‌ترین ابزارهای داده‌کاوی به ترتیب پرترفدار بودن، زبان برنامه‌نویسی آر (Venables et al, 2016)، زبان برنامه‌نویسی پایتون (Miller, Bryce, 2016)، Clementine که نسخه ۱۳ آن با نام SPSS Modeler نامیده می‌شود، نرم‌افزار وکا و نرم افزار متلب می‌باشد.

## مفاهیم تعیین کننده در داده‌کاوی

هم پیوندی (Bagging): هم پیوندی قابلیتی برای یافتن روابط ناشناخته موجود در اطلاعات است. این روابط مواردی از قبیل اینکه حضور مجموعه‌ای از مقولات اشاره به این دارند که مجموعه مقولات دیگری نیز احتمالاً وجود دارند را شامل می‌شود. این قابلیت اساساً روشی است برای اینکه کشف کنیم چه مقولاتی به هم می‌خورند. از آن با عنوان تحلیل سبد بازار یا گروه‌بندی خوشه‌بندی نیز یاد می‌شود.

طبقه‌بندی (Boosting): طبقه‌بندی در واقع ارزشیابی ویژگی‌های مجموعه‌ای از داده‌ها و سپس اختصاص دادن آن‌ها به مجموعه‌ای از گروه‌های از پیش تعریف شده است. این متداول‌ترین قابلیت داده‌کاوی می‌باشد. داده‌کاوی را می‌توان با استفاده از داده‌های تاریخی برای تولید یک مدل یا نمایی از یک گروه بر اساس ویژگی‌های داده‌ها به کار برد. سپس می‌توان از این مدل

تعریف شده برای طبقه بندی مجموعه داده‌های جدید استفاده کرد. همچنین می‌توان با تعیین نمایی که با آن سازگار است برای پیش بینی‌های آتی از آن بهره گرفت.

الگوهای ترتیبی: قابلیت‌های ترتیبی هم مانند قابلیت‌های هم پیوندی این خاصیت را دارند که می‌توانند وقایع را به هم مرتبط کنند. این کار در هم پیوندی سنتی یا تحلیل سبب بازار مجموعه‌ای از مقولات را به عنوان مقولات پشت سر هم ارزیابی می‌کنند و از ابزارهایی مثل سری‌های زمانی برای تعیین ترتیب بهره می‌برند. الگوهای ترتیبی علاوه بر آن، این قابلیت جدید را هم دارند که می‌توانند فاصله زمانی بین دو واقعه را تخمین بزنند. برای مثال این قابلیت امکان نتیجه‌گیری‌هایی از قبیل اینکه "هشتاد درصد افرادی که کامپیوتر می‌خرند ظرف مدت یک سال چاپگر هم خواهند خرید." را مهیا می‌نماید. به این ترتیب شناسایی نوعی از خریدهای مقدماتی که پتانسیل خریدهای بعدی را در آینده تعیین می‌کنند، عملی می‌شود. در نتیجه از چنین تحلیل‌هایی به شدت در تبلیغات فروش استفاده می‌گردد.

خوشه بندی (clustering): قابلیت خوشه بندی وظیفه تقسیم یک گروه ناهمجنس را در چندین زیر گروه بر عهده دارد. این فرایند یک تفاوت اساسی با طبقه بندی دارد. زیرا در این مدل هیچ گونه الگوی آموزشی نداریم. خوشه بندی به طور خودکار ویژگی‌های متمایز کننده زیر گروه‌ها را تعریف می‌کند و زیر گروه‌ها را سازماندهی می‌نماید. و به عنوان نوعی قابلیت داده کاوی غیر مستقیم مطرح است. این ابزارها پایگاه داده را بر اساس ویژگی‌های داده‌ها به چندین بخش تقسیم می‌کنند و گروه‌هایی از رکوردها را به وجود می‌آورند که نمایانگر یا صاحب صفت خاصی هستند. الگوهای به دست آمده در ذات پایگاه داده نهاده شده هستند و نشانگر بعضی اطلاعات غیر منتظره و در عین حال ارزشمند شرکتی باشند.

به منظور طبقه بندی کاربردهای داده کاوی در شرکت‌های مخابرات، رویکردهای مختلفی وجود دارد که از آن جمله می‌توان به موارد زیر اشاره کرد:

طبقه بندی بر اساس روش‌های داده کاوی: این طبقه بندی شامل خوشه بندی، رگرسیون، دسته بندی، کشف الگو و ... می‌باشد. در این نوع طبقه بندی، هدف ارزیابی و بهبود تکنیک‌های داده کاوی بکار رفته و بکارگیری روش‌های جدید برای حل مسائل است.

طبقه بندی بر اساس حوزه کاری: در این نوع طبقه بندی، هدف بهبود و ارائه راهکارهای جدید داده کاوی در حوزه‌های موجود در سازمان‌های مخابرات است. بعنوان مثال تلفن ثابت، تلفن همراه، دیتا و ADSL.

طبقه بندی بر اساس هدف: در این نوع طبقه بندی، هدف شناسایی حوزه‌های جدید داده کاوی در مخابرات با محوریت جذب و حفظ مشتری است که از این گروه می‌توان به بازاریابی هدفمند، افزایش سود، افزایش رضایت مشتری اشاره نمود.

بر اساس مطالعه و بررسی انجام شده، می‌توان گفت طبقه بندی جامعی از کارهای انجام شده در صنعت مخابرات انجام نشده است. از آنجاییکه هدف اصلی شرکت‌های مخابراتی، افزایش سود و رضایت مشتری از طریق ارائه خدمات متناسب به مشتریان می‌باشد، رویکرد این تحقیق برای طبقه بندی کاربردهای داده کاوی، مبتنی بر رفتار مشتریان است. برای این منظور با استفاده از مدل پیشنهادی LRFM که در بخش ۳ توضیح داده شده است خوشه بندی را انجام می‌دهیم.

در ادامه مقاله به تحقیقات پیش در خصوص خوشه بندی مشتریان می‌پردازیم در بخش بعدی مدل پیشنهادی که به دسته بندی مشتریان بر اساس توسعه مدل RFM به LRFM را توضیح می‌دهیم و به ارزیابی و اجرای این الگوریتم در محیط کلمنتاین پرداخته و در بخش آخر به نتیجه‌گیری نهایی در خصوص این تحقیق خواهیم پرداخت.

### تحقیقات پیشین در زمینه روش‌های داده کاوی

در مطالعه‌ی صورت گرفته در (Clemens et al, 2010) نویسندگان مدلی را برای تجزیه و تحلیل رفتار خرید مشتریان در سوپر مارکت شبیه سازی کردند. این مدل شبیه سازی، دانشی را در مورد آنچه که مشتری داخل فروشگاه انجام می‌دهد را از قبیل نحوه حرکت در فروشگاه، چیدمان محصولات، خرید و یا نخریدن محصولات را نشان می‌دهد.

در (Wei, 2011) نویسنده از مدل RFM برای طبقه بندی مشتریان خاص خود استفاده می کند. این مطالعه در طرح برنامه های بازاریابی برای مشتریان مختلف بسیار مفید است.

در مقاله ی (Divya et al, 2013) نویسندگان از تجزیه و تحلیل مدل RFM برای بدست آوردن ارزش مشتریان در آینده استفاده کرده اند. تجزیه و تحلیل RFM کمک به بهبود رابطه با مشتریان می کند.

در تحقیق (Raorane and.Kulkarni, 2011) نویسندگان به تحلیل مشتریان قابل اعتماد در صنعت سوپر مارکت ها پرداختند. در این تحقیق به تحلیل محصولاتی که توسط مشتریان با هم خریداری می شود پرداخته شده است.

در مطالعه صورت گرفته شده در (MARGIANTI et al, 2016) یک فرایند داده کاوی با استفاده از مدل RFM بر داده های ۱۰۰۰ مشتری انجام گرفت. به کمک نرم افزار متلب و تحلیل تراکنش های داده ای مشتریان ۵۱ خوشه تولید و سپس مشتریان به هشت دسته تقسیم شدند.

در (Aniruddha and Devansh, 2016) پیاده سازی مدل جدیدی بر مبنای مدل RFM ارائه شد. در این مدل وفاداری مشتری نیز به عنوان فاکتور جدیدی در مدل معرفی می شود. این مدل شامل سه فاکتور تازگی خرید، تکرار یا فرکانس و وفاداری مشتری می باشد. هدف اصلی این مطالعه تقسیم بندی مشتریان و ارزیابی آن ها بر اساس تاریخچه گذشته آنان می باشد. دو عیب عمده مدل RFM بطور مستمر قراردادن مشتریانی با امتیاز بالای RFM (مشتریانی که بصورت مکرر مورد هدف بودند همین سبب آزار و اذیت آن ها می شد) و نقطه ضعف دیگر مدل غفلت از مشتریانی است که امتیاز کمتری دارند. یک شرکت می تواند اعتبار خود را با در نظر گرفتن مشتریان گذشته خود افزایش دهد. در این مدل مشتریان قدیمی بطور مناسب و منظم مورد بررسی قرار می گیرند.

در (Aarati et al, 2016) مدل جدید TRFM برای تحلیلی و آنالیز داده های مرتبط با مشتریان معرفی شد. در این مطالعه به کمک Eclipse Java EE IDE به پیاده سازی مدل پرداخته شده است و داده های ۸۴۰۰ مشتری سوپر مارکت مورد بررسی قرار گرفته است. در مدل RFM سنتی به تازگی، تکرار و مبلغ خرید مشتریان توجه می شود. در حالی که در مدل TRFM علاوه بر فاکتورهای فوق، به مدت زمان سپری شده توسط مشتری در سوپر مارکت نیز توجه می شود. مشتریان با بالاترین نمره TRFM بهترین مشتریان هستند. آن ها مشتریانی هستند که زمان کمتری را صرف خرید در سوپر مارکت می کنند و به نسبت مشتریانی که وقت بیشتری را صرف می کنند بیشتر مورد پسند واقع می شوند.

در مطالعه صورت گرفته شده در (hamdi and zamiri, 2016) هدف ارائه چارچوبی برای تقسیم بندی مشتریان بیمه خودرو شرکت بیمه پاسارگاد براساس فاکتورهای مؤثر بر ارزش طول عمر مشتریان می باشد. برای این منظور یک سری از معاملات مرتبط با ۳۸۴ مشتری شرکت بیمه پاسارگاد در بهار ۲۰۱۵ بصورت تصادفی در نظر گرفته شد. اطلاعات معامله شامل تازگی خرید مشتریان، تکرار یا فرکانس تمدید بیمه نامه در مدت شش سال و مقدار پولی که مشتری در آخرین قرارداد بیمه خود پرداخت کرده است می باشد. با توجه به مدل RFM و تحلیل این داده ها، مشتریان به ۴ دسته طلایی، با ارزش وفادار، ثابت و فوق العاده با ارزش تقسیم شدند.

در (Oluwafemi, 2016) به تجزیه تحلیل اطلاعات ۱۸۰ مشتری که برای صرف نهار به یک رستوران مراجعه کرده اند به کمک مدل RFM پرداخته شد. از ۱۸۰ مشتری مورد بررسی قرار گرفته شده تعداد ۱۰۰ نفر آن ها در سیستم امتیاز بندی دارای امتیاز بالایی بودند. این نتیجه به مدیریت رستوران کمک می کند برنامه های مؤثری را برای تبلیغات خود در آینده داشته باشند.

در مطالعه صورت گرفته در (Mohammadian and Makhani, 2016) از مدل RFM برای تحلیل و طبقه بندی مشتریان استفاده شده است. مشتریان طی این تجزیه تحلیل به هشت دسته منطقی مجزا تقسیم شدند. نتایج حاصل کمک می کند تا برای بهبود فروش، بازاریابی و تصمیم گیری در محیط های خرده فروشی بصورت رقابتی تصمیمات بهتری گرفته شود.

بدحسابی و خوش حسابی مشتریان به خصوص برای برخی شرکت‌های مخابراتی که ابتدا سرویس ارائه می‌دهند و سپس صورتحساب صادر می‌کنند مسئله مهمی است و اغلب تاثیر زیادی بر درآمدها و مطالبات آن‌ها دارد. در همین راستا، در (Amal et al, 2014) نویسنده چند تکنیک داده کاوی را در خصوص شناسایی مشتریان بکار گرفته است و نتایج تحلیلی آن را ارائه کرده است.

در پژوهش دیگری (فروش، ۱۳۸۷) که با هدف شناسایی مشترکین بد حساب و اعلام زمان مناسب برای عکس العمل انجام شده است، از داده‌های مرکز مخابرات تهران استفاده شده و با بکارگیری تکنیک‌های داده کاوی، یک مدل داده کاوی برای شناسایی مشترکین بدحساب شرکت‌های مخابرات ایران ارائه شده است.

در پژوهش دیگری که در شرکت مخابرات خراسان جنوبی انجام شده است (صداقتی مختاری و همکاران، ۱۳۹۱) یک سیستم نرم افزاری هوشمند مبتنی بر تکنیک‌های داده کاوی برای طبقه بندی مشترکین این شرکت طراحی و پیاده سازی شده است. در این پروژه، داده‌های مربوط به مشترکین تلفن ثابت، تلفن همراه، مشترکین دیتا و ADSL مورد کاوش قرار گرفته اند.

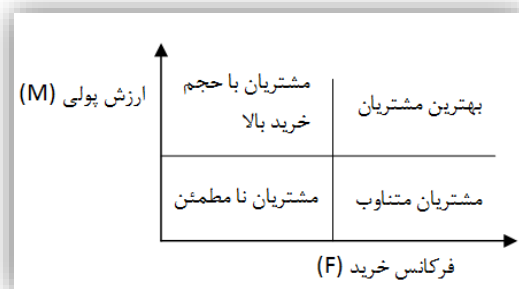
در تحقیق دیگری (ایسوند و حسینی، ۱۳۹۳)، که براساس داده‌های یکی از مراکز مخابرات کشور جمع آوری شده است، چارچوبی برای طبقه‌بندی مشترکین مخابرات تلفن ثابت با به کارگیری مدل RFM ارائه شده که بر اساس مدل بدهکاری، مشتری بر مبنای وضعیت کارکرد از دو بعد ارزش و ضد ارزش سنجیده می‌شود. در انتها نیز بر اساس بخش بندی انجام گرفته با کمک تکنیک‌های داده کاوی، راهبردهایی برای برخورد با هر یک از گروه‌های مشتری، ارائه شده است.

در پایان نامه (بهادری، ۱۳۸۸) سعی بر این بوده است تا با استفاده از داده‌های مصرف مشترکین و اطلاعات حاصل از پروفایل مشترکین، به نیازهای هر بخش پی برده و از این دانش ایجاد شده در حوزه‌های مختلفی چون تولید محصول، مهندسی ترافیک و مدیریت ارتباط با مشترکین استفاده شود. در این تحقیق مدلی به منظور خوشه بندی مشترکین ADSL پیشنهاد شده است که مشتریان را به پنج قسمت تقسیم کرده است.

### توسعه مدل RFM به LRFM

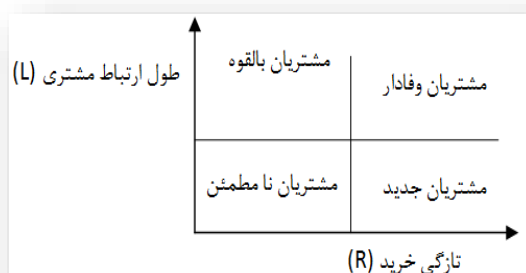
مدل LRFM روشی برای خوشه‌بندی مشتریان است که در مدیریت ارتباط با مشتری استفاده می‌کنیم. در این مدل مشتریان براساس چهار ویژگی طول ارتباط مشتری، تازگی خرید، تناوب خرید و ارزش مالی خرید دسته‌بندی می‌شوند. براساس حرف اول هر یک از این چهار ویژگی واژه LRFM نوآوری شده است. در این تحقیق با استفاده از مدل مذکور اقدام به آموزش خوشه‌بندی با نرم‌افزار کلمنتاین می‌پردازیم.

بر طبق (Reinartz and Kumar, 2002 - Chang and Tsay, 2004) مدل RFM نمی‌تواند مشتریان دارای ارتباط بلند مدت و مشتریان دارای ارتباط کوتاه مدت با سازمان را مشخص نماید. آن‌ها در تحقیق خود ایده طول ارتباط مشتری را پیشنهاد می‌دهند و به بررسی تاثیر آن بر وفاداری و سود آوری مشتری می‌پردازند. آن‌ها بیان می‌کنند که افزایش طول ارتباط با مشتری، وفاداری مشتری را بهبود خواهد بخشید. و این متغیر را که نشان دهنده فاصله زمانی بین اولین و آخرین خرید مشتری در بازه مورد مشاهده است تعریف کرده‌اند. مدل RFM مشتریانی را که به تازگی ارزش مالی بالایی برای شرکت ایجاد کرده و در کوتاه مدت دارای تناوب خرید بیش تر از متوسط تناوب خرید در بین مشتریان را به عنوان مشتریان با ارزش انتخاب کرده در حالی که عامل طول ارتباط با شرکت نادیده گرفته شده است. بنابراین بعد طول ارتباط مشتری Length(L) به مدل RFM اضافه می‌شود. طول ارتباط مشتری با سازمان، نشان‌دهنده مدت زمانی است که یک مشتری ارتباط خود را با سازمان آغاز کرده است. بر طبق مطالعات انجام شده در [۲۰] با اضافه شدن شاخص L (طول ارتباط مشتری) زمینه جهت تحلیل دقیق تر مشتریان فراهم می‌گردد. آن‌ها یک ماتریس با عنوان ماتریس ارزش به ابعاد فرکانس خرید (F) و ارزش پولی (M) پیشنهاد می‌کنند که در شکل ۱ به تصویر کشیده شده است.



شکل ۱: ماتریس ارزش [۲۰]

همچنین ادعا می‌کند که ارتباط طولانی‌تر مشتری، وفاداری بالاتر و کوتاه‌تر بودن زمان تعامل اخیر، بالاتر بودن وفاداری مشتری را نشان می‌دهد. دو شاخص دیگر، طول ارتباط مشتری (L) و زمان معامله اخیر (R) به عنوان ماتریس وفاداری مشتری تعریف می‌شود. این ماتریس در شکل ۲ به تصویر کشیده شده است.



شکل ۲: ماتریس وفاداری مشتری [۲۰]

در این تحقیق همانند مدل ارائه شده در (مصلحی و همکاران، ۱۳۹۳) به بخش‌بندی و محاسبه ارزش دوره عمر مشتریان و شناسایی مشتریان ارزشمند استفاده می‌پردازیم.

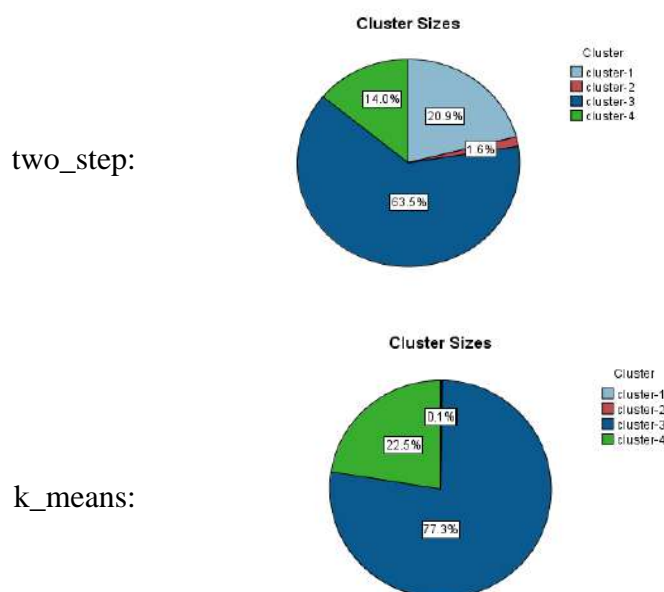
### ارزیابی و اجرا:

در این بخش بر اساس داده‌های سیستم خدمات اینترنتی مخابرات خراسان جنوبی به پیاده‌سازی مدل می‌پردازیم. بخش‌بندی مشتریان موجب شناخت صحیح مشتریان می‌گردد تا در تعریف و تصمیم‌گیری طرح‌های تبلیغاتی تخصصی‌تر مورد استفاده قرار گیرد. برای بخش بندی مشتریان از خوشه‌بندی که یکی از روش‌های داده‌کاوی است استفاده می‌کنیم و براساس این، مشتریان در غالب مشتریان وفادار، مشتریان بالقوه، مشتریان از دست رفته، مشتریان جدید و مشتریان با مصرف زیاد خدمات تقسیم می‌شوند. پایگاه داده سوابق که شامل جدول سوابق خرید ۹۹۹ مشتری در طی یکسال، شامل ۱۰۰۰ رکورد و ۴ فیلد شامل طول ارتباط مشتری و تازگی آخرین خرید و تعداد تکرار خرید و ارزش پولی خرید برای خوشه بندی مورد استفاده قرار می‌گیرد. بخشی از جدول سوابق مشتریان در شکل ۳ به تصویر کشیده شده است.

	L-N	R-N	F-N	M-N	
1		0.758	0.585	0.000	1.000
2		0.150	0.944	0.002	0.478
3		0.999	0.958	0.003	0.349
4		0.150	0.850	0.000	0.239
5		0.928	0.977	0.003	0.232
6		0.996	0.268	0.001	0.217
7		1.000	0.838	0.001	0.207
8		0.056	0.944	0.000	0.143
9		0.150	0.850	0.000	0.130
10		0.795	0.912	0.002	0.118
11		0.877	0.252	0.000	0.115
12		0.718	0.962	0.004	0.113
13		0.054	0.980	0.001	0.104
14		0.470	0.988	0.000	0.102
15		0.457	0.543	0.000	0.100
16		0.596	0.841	0.004	0.099
17		1.000	0.696	0.007	0.098
18		0.996	0.838	0.003	0.097
19		0.996	0.843	0.003	0.091
20		0.996	0.963	0.005	0.087

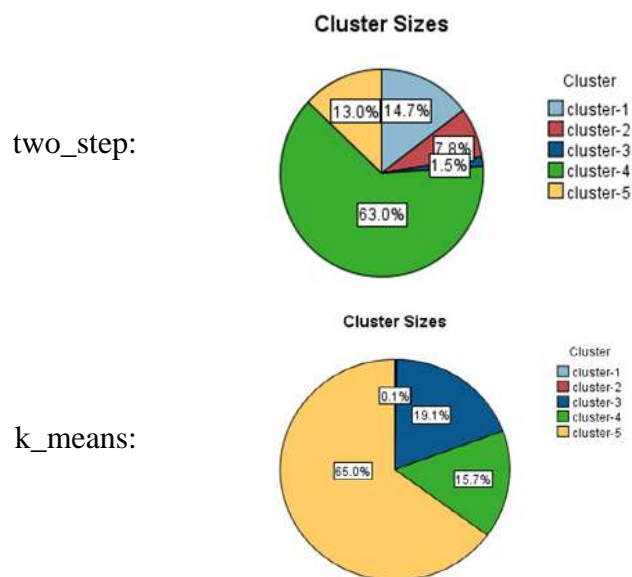
شکل ۳: سوابق خرید مشتریان در پایگاه داده

برای تعیین ارزش مشتریان از مدل LRFM استفاده می‌شود در این مرحله مشتریان براساس متغیرهای L,R,F,M نرمال شده‌اند و با استفاده از دو الگوریتم TwoStep و k\_Means (Mooi and Sarstedt, , 2011) با استفاده از نرم افزار کلمنتاین خوشه بندی شده‌اند. نتایج خوشه بندی در شکل های ۴ تا ۹ در ادامه آمده است.



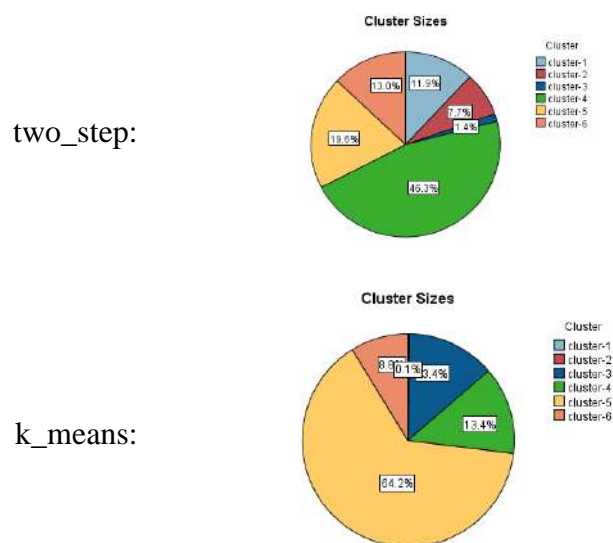
شکل ۴: خوشه بندی two\_step و k\_means با ۴ خوشه

همانطور که در شکل ۴ مشاهده می‌کنید خوشه های ۱ و ۲ در الگوریتم K\_means دارای جمعیت بسیار کمی می باشد. اما در الگوریتم Two step این دو خوشه بخشی از جمعیت را دارا می باشند. شکل ۵ مقایسه ای بین خوشه های تولید شده توسط دو الگوریتم مذکور را نشان می‌دهد. در این شکل با وجود خوشه‌بندی به تعداد ۵ خوشه در روش K\_means جمعیت خوشه های ۱ و ۲ غیر قابل توجه هستند.



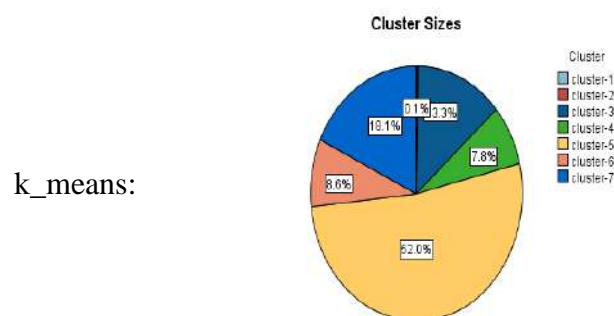
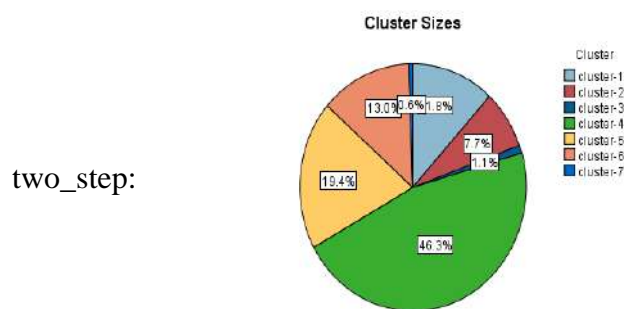
شکل ۵: خوشه بندی two\_step و k\_means با ۵ خوشه

شکل ۶ مقایسه نتایج حاصل از خوشه بندی به شش خوشه را توسط دو الگوریتم نشان می‌دهد. در این شکل نیز شاهد آن هستیم که جمعیت در الگوریتم K\_means بین خوشه‌ها توزیع نشده است.



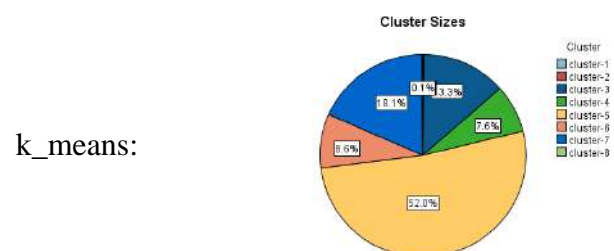
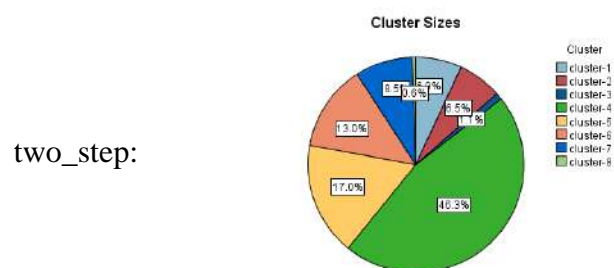
شکل ۶: خوشه بندی two\_step و k\_means با ۶ خوشه

شکل ۷ خوشه بندی به هفت خوشه را توسط دو الگوریتم نشان می‌دهد. که با توجه به نتایج به تصویر کشیده شده تفاوت دو الگوریتم قابل مشاهده است.



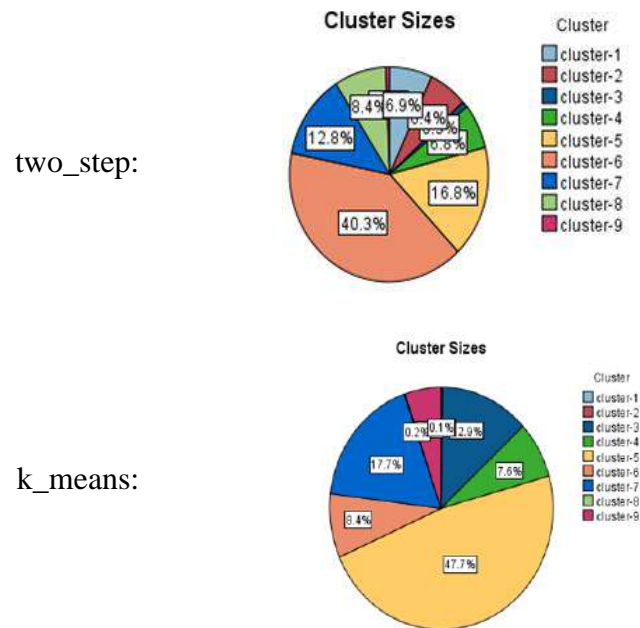
شکل ۷: خوشه بندی two\_step و k\_means با ۷ خوشه

در شکل ۸ نتایج خوشه بندی به هشت خوشه نشان داده شده است. همانطور که مشاهده می شود الگوریتم two\_step تعداد خوشه هایی با جمعیت های توزیع شده تر را دارا می باشد.



شکل ۸: خوشه بندی two\_step و k\_means با ۸ خوشه

در شکل ۹ نیز خوشه بندی مشتریان به نه خوشه صورت گرفته است. همانطور که مشاهده می کنید در الگوریتم K\_means تعداد زیادی از مشتریان در یک خوشه و بقیه در هشت خوشه باقیمانده توزیع شده اند.



شکل ۹: خوشه بندی two\_step و k\_means با ۹ خوشه

پس از اتمام فاز خوشه بندی به ارزیابی خوشه های بدست آمده باید پرداخت. در این مرحله خوشه بندی انجام گرفته و با استفاده از شاخص سنجش کیفیت دیویس \_ بولین همانند ارزیابی صورت گرفته در (Sandhya and Surya, 2015) به ارزیابی خوشه ها می پردازیم. در شاخص سنجش کیفیت دیویس - بولین کمترین شباهت بین خوشه ای و بیشترین شباهت درون خوشه ای مد نظر می باشد. نتایج شاخص دیویس - بولین در جدول ۱ بصورت خلاصه آمده است.

جدول ۱: نتایج شاخص دیویس

DB index	تعداد خوشه ها
0/178	4
0/0678	5
0/0435	6
0/0555	7
0/0486	8
0/0432	9

با توجه به محاسبه شاخص دیویس که حاصل تناسب دو شاخص تراکم و جدایی می باشد و بستگی به تعداد خوشه ها دارد می توان از جدول ۱ نتیجه گرفت که اگر تعداد خوشه ها ۹ باشد مقدار شاخص دیویس کمتر و در نتیجه خوشه بندی مناسب تری داریم. برای تشخیص اینکه کدام روش خوشه بندی مناسب تر می باشد می توان از روی نمودار آن ها نتیجه گرفت که الگوریتم k\_Means نسبت به الگوریتم TwoStep ناکارآمدتر می باشد زیرا بخش اعظم داده ها در یک خوشه قرار گرفته اند. بنابراین الگوریتم TwoStep با تعداد ۹ خوشه مناسب تر می باشد.

در جدول ۲ تعداد داده هر خوشه و مختصات نقاط مرکزی ۹ خوشه مشخص شده اند که با توجه به این اطلاعات به تحلیل خوشه ها می پردازیم.

جدول ۲: نتایج خوشه بندی Twostep

Cluster	cluster 4	cluster 5	cluster 7	cluster 8	cluster 1	cluster 4	cluster 2	cluster 3	cluster 6
Label									
Percentage									
Size	45.7% (407)	16.8% (158)	12.9% (120)	6.4% (59)	8.8% (80)	6.8% (62)	6.4% (59)	0.3% (3)	9.0% (84)
Weight	LN 3.92	LN 0.00	LN 0.01	LN 0.01	LN 0.04	LN 0.01	LN 0.01	LN 0.00	LN 0.00
	FN 0.01	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00
	FN 0.01	FN 0.01	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00	FN 0.00
	MN 0.00	MN 0.00	MN 0.04	MN 0.00	MN 0.00	MN 0.00	MN 0.00	MN 0.00	MN 0.00

با فرض اینکه نقاط مرکزی هر خوشه نماینده کل نقاط خوشه هستند ویژگی مشتریان تحلیل می‌شود در ادامه ویژگی های هر خوشه ارائه می‌شود.

خوشه ۱: شامل ۶۹ مشتری بوده و ۶/۹٪ از مشتریان کل را تشکیل می‌دهد و مختصات نقاط مرکزی نشان می‌دهد که طول ارتباط مشتری و تازگی آخرین خرید نسبت به میانگین‌شان کمتر بوده پس این خوشه جزوه مشتریان جدید بوده است.

خوشه ۲: شامل ۶۴ نفر مشتری بوده و ۶/۴٪ از مشتریان کل را تشکیل می‌دهد و مختصات نقاط مرکزی نشان می‌دهد که طول ارتباط مشتری نسبت به خوشه ۱ بیشتر می‌باشد ولی باز هم نسبت به میانگین کمتر بوده و تازگی آخرین خریدشان نسبت به خوشه ۱ کمتر بوده پس این خوشه هم جزوه مشتریان جدید هستند که از آخرین خریدشان زمان کمتری نسبت به خوشه ۱ می‌گذرد.

خوشه ۳: افراد این خوشه رفتارشان نسبت به دو خوشه ۱ و ۲ متفاوت می‌باشد افراد این خوشه فقط ۹ نفر می‌باشند و ۰/۹٪ از افراد کل را تشکیل می‌دهند. مختصات نقاط مرکزی‌شان نشان می‌دهد که طول ارتباط مشتری مانند دو خوشه ۱ و ۲ کم می‌باشد اما از آخرین خرید آنان مدت زیادی گذشته بطوری که می‌توان آنان را در دسته مشتریان از دست رفته قرار داد.

خوشه ۴: این خوشه شامل ۶۸ نفر می‌باشد یعنی ۶/۸٪ از مشتریان کل می‌باشند. با توجه به مختصات نقاط مرکزی می‌توان نتیجه گرفت که طول ارتباط مشتری زیاد بوده یعنی مدت زمانی زیادی می‌گذرد که مشتری از این سیستم استفاده می‌کند و تازگی آخرین خرید آن هم کم بوده یعنی جدیداً خریدی را انجام داده و با توجه به اینکه فرکانس خرید آن پایین تر از میانگین بوده و مبلغ خرید آن زیاد بوده می‌توان نتیجه گرفت که این دسته از مشتریان، مشتریان وفادار هستند که مدت طولانی از سیستم استفاده می‌کنند و تعداد دفعات خرید آنان کم ولی مبالغ خرید آنان زیاد می‌باشد.

خوشه ۵: این خوشه شامل ۱۶۸ نفر می‌باشد که ۱۶/۸٪ از مشتریان را تشکیل می‌دهد این خوشه مانند خوشه ۳ می‌باشد با این تفاوت که طول ارتباط مشتری نسبت به آن بیشتر بوده و فرکانس آن نیز نسبت به میانگین زیاد می‌باشد اما این مشتریان در دسته مشتریان از دست رفته قرار می‌گیرند.

خوشه ۶: این خوشه که بیشترین تعداد مشتری را دارد شامل ۴۰۳ نفر می‌باشد که ۴۰/۳٪ از مشتریان کل را تشکیل می‌دهد با توجه به مختصات مرکزی آن می‌توان نتیجه گرفت که طول ارتباط مشتری زیاد بوده و تازگی آخرین خرید آن زیاد می‌باشد یعنی مدت طولانی از آخرین خرید آن‌ها می‌گذرد و فرکانس و مبالغ خرید آن‌ها زیاد می‌باشد این دسته از مشتریان در دسته مشتریان بالقوه قرار می‌گیرند.

خوشه ۷: این خوشه شامل ۱۲۸ نفر می‌باشد یعنی ۱۲/۸٪ از مشتریان کل را تشکیل می‌دهد این خوشه مانند خوشه ۴ می‌باشد با این تفاوت که تازگی آخرین خرید آن نسبت به خوشه ۴ کمتر بوده یعنی از آخرین خرید آن زمان زیادی نمی‌گذرد و مبالغ خرید آن نیز نسبت به آن خوشه بیشتر می‌باشد و در دسته مشتریان وفادار قرار می‌گیرند.

خوشه ۸: این خوشه شامل ۸۴ نفر می‌باشد که ۸/۴٪ از مشتریان را تشکیل می‌دهد این خوشه نیز مانند خوشه ۳ و ۵ می‌باشد با این تفاوت که طول ارتباط مشتری آن نسبت به دو خوشه دیگر کمتر می‌باشد و در دسته مشتریان از دست رفته قرار می‌گیرند. خوشه ۹: این خوشه که کمترین افراد را دارد شامل ۶ نفر می‌باشد یعنی ۰/۶٪ از مشتریان را تشکیل می‌دهد این خوشه نیز مانند خوشه‌های ۳ و ۵ و ۷ می‌باشد با این تفاوت که طول ارتباط مشتری آن خیلی کمتر از بقیه خوشه‌ها می‌باشد و فرکانس و مبالغ خرید آن زیاد می‌باشد بطوری که فرکانس خرید آن از سایر خوشه‌ها خیلی بیشتر است. می‌توان نتیجه گرفت که مشتریان این خوشه در مدت خیلی کوتاه تعداد دفعات خرید زیادی داشته‌اند و از خرید آنان مدت زیادی می‌گذرد و در دسته مشتریان از دست رفته قرار می‌گیرند.

### نتیجه‌گیری و پیشنهادات

با توجه به آنچه گفته شد می‌توان مشتریان را در پنج دسته تقسیم کرد:

**مشتریان وفادار:** مشتریان خوشه ۴ مشتریان وفادار می‌باشند که همگی طول ارتباط مشتری آن‌ها از میانگین بیشتر و شاخص تازگی آن‌ها از میانگین کمتر بوده و فرکانس خرید و مبالغ خرید آن‌ها حالت‌های مختلفی دارد این مشتریان مدت طولانی از این سیستم استفاده کرده‌اند و تاکنون هم به همکاری خود با مخابرات ادامه داده‌اند. برای این دسته از مشتریان می‌توان سیاستی در نظر گرفت که از همکاریشان با این مجموعه تشکر شود.

**مشتریان بالقوه:** مشتریان خوشه ۶ از مشتریان بالقوه هستند که همگی طول ارتباط و تازگی خرید آن‌ها از میانگین بیشتر است و فرکانس خرید و مبالغ آن‌ها حالت‌های مختلف دارد این مشتریان مدت طولانی از این سیستم استفاده کرده‌اند ولی اخیراً تازگی خرید آن‌ها کاهش یافته این نشان می‌دهد که این ممکن است تعدادی از این مشتریان از جزوه مشتریان از دست رفته قرار بگیرند برای حفظشان باید سیاستی اندیشد که رضایتشان افزایش یابد و خریدشان نیز افزایش پیدا کند.

**مشتریان جدید:** مشتریان خوشه ۱ و ۲ مشتریان جدید می‌باشند. که طول ارتباط مشتری و تازگی خرید آن‌ها از میانگین کمتر می‌باشد و فرکانس خرید و مبالغ خرید آن‌ها حالت‌های مختلفی دارد این مشتریان سابقه ارتباط طولانی مدت با آن‌ها وجود ندارد برای این دسته از مشتریان باید استراتژی و برنامه‌ای را ریخت که همکاریشان حفظ شود زیرا حفظ مشتری از جذب مشتری اهمیت بیشتری دارد و کار مشکل‌تری می‌باشد.

**مشتریان از دست رفته:** مشتریان خوشه ۳ و ۵ و ۸ و ۹ مشتریان از دست رفته می‌باشند که این گروه از مشتریان ارتباط طولانی نداشته و اخیراً هم خرید نداشته‌اند و فرکانس خرید و مبالغ خرید آن‌ها حالت‌های مختلفی دارد باید استراتژی و برنامه‌ای طرح کرد که علت اینکه این مشتریان از دست رفته‌اند مشخص شود و کاری کرد که مشتریان پر ارزش از سیستم خارج نشوند.

**مشتریان پر مصرف:** این مشتریان فرکانس خرید و مبالغ خریدشان از میانگین کمتر می‌باشد و طول ارتباط مشتری و تازگی خریدشان حالت‌های مختلفی دارد خوشبختانه از این تعداد ۱۰۰۰ نفر مشتری هیچکدام در این گروه قرار نگرفته‌اند در صورت وجود افرادی در این گروه باید سیاستی طرح ریزی کرد که بعد از اتمام دوره تمدید قرار دارد با این افراد صورت نگیرد. در این پژوهش توانستیم مشتریان مخابرات را خوشه بندی کنیم و تعداد اعضای هر خوشه را مشخص کردیم ولی این پژوهش می‌تواند گسترده‌تر شود مثلاً می‌توان مشتریان را به دسته‌های بیشتر تقسیم کرد و معیارهای دیگر مثل شغل و ناحیه جغرافیایی را در ملاک‌های تقسیم بندی قرار داد. با توجه به آنچه گفته شد مدل LRFM برای ارزش گذاری مشتریان مخابرات مناسب می‌باشد از این مدل می‌توان در زمینه‌های دیگر همچون بانک استفاده کرد.

ایسوند، فروغ، حسینی، منیره، (۱۳۹۳)، " رویکردی نوین در بخش بندی مشترکین مخابرات تلفن ثابت با به کارگیری مدل RFM " ، مجله ی تخصصی مهندسی صنایع، دوره ۴، شماره ۱

بهادری، سعید (۱۳۸۸)، "بخش بندی مشترکین خدمات ADSL بر مبنای داده کاوی"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس.

صداقتی مختاری، محمود رضا، رضوانی خراشادیزاده، رامین، احمدی درمیان، احسان، اکبری، محسن (۱۳۹۱)، "طراحی و پیاده سازی یک سیستم نرم افزاری هوشمند مبتنی بر تکنیک های داده کاوی برای طبقه بندی مشترکین شرکت مخابرات خراسان جنوبی"، ششمین کنفرانس داده کاوی ایران

کاتلر، فیلیپ، آرمسترانگ، گری (۱۳۸۶ اصول بازاریابی). (بهمن فروزنده، مترجم). اصفهان، نشر آموخته. ۱۵ فروروش، حمید، (۱۳۸۷)، " ارائه یک مدل داده کاوی برای شناسایی مشترکین بدحساب مورد کاوی شرکت های مخابرات ایران"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس

مصلحی، سیده نیره، کفاش پور، آذر، ناجی عظیمی، زهرا، (۱۳۹۳)، "استفاده از مدل LRFM برای بخش بندی مشتریان براساس ارزش چرخه عمر آنها"، پژوهش های مدیریت عمومی

Aarati Joshi, Dr. Vipul Vekariya, Sagar Patel,(2016), "Design & Analysis of Purchasing Behaviour of Customers in Supermarkets using TRFM Model of Data Mining", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 4

Amal M. Almana, Mehmet Sabih Aksoy, Rasheed Alzahrani,(2014), "A Survey On Data Mining Techniques In Customer Churn Analysis For Telecom Industry", Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 5( Version 6), pp.165-171

Aniruddha Kamath, Devansh Dave, (2016), "DA-RFM Model for Measuring Prospective Customer Value", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 6

Chang H. H. and Tsay S. F. (2004), " Integrating of SOM and K-Mean in data mining clustering: an empirical study of CRM and profitability evaluation", J. Inf. Manage, 11 (4): 161-203.

Clemens Schwenke, Volodymyr Vasyutynskyy and Klaus Kabitzsch, (2010), "Simulation and Analysis of Buying Behavior in Supermarkets", IEEE Conference on Emerging Technology and Factory Automation (ETFA), pp. 1-4

Divya D. Nimbalkar and Asst Prof. Paul, (2013), "Data mining using RFM Analysis", International Journal of Scientific & Engineering Research, Vol. 4, Issue 12, pp.940-943

hamdi, Karim , zamiri, Ali , (2016), "identifying and segmenting customers of Pasargad insurance company through RFM model (RFM)", International Business Management, 10(18): 4209-4214, ISSN: 1993-5250

Mooi E and Sarstedt M., (2011), "A Concise Guide to Market Research", DOI 10.1007/978-3-642-12541-6\_9, # Springer-Verlag Berlin Heidelberg 2011

MARGIANTI, REFIANTI, MUTIARA, NUZULINA, (2016), "AFFINITY PROPAGATION AND RFM-MODEL FOR CRM'S DATA ANALYSIS", Journal of Theoretical and Applied Information Technology. Vol.84. No.2

Mohammadian Mahmoud, Makhani, Iman, (2016), "RFM-Based customer segmentation as an elaborative analytical tool for enriching the creation of sales and trade marketing strategies", International Academic Journal of Accounting and Financial Management Vol. 3, No. 6, pp. 21-35. ISSN 2454-2210

Miller Preston, Bryce Chapin, Learning Python for Forensics, May 2016

Oluwafemi Elijah OJO, (2016), "Analysis of Purchasing Behaviour of Customers in Restaurants Using RFM Model of Data Mining", International Journal of Research, p-ISSN: 2348-6848 e-ISSN: 2348-795X, Volume 03 Issue 13

Reinartz, Werner J., & Kumar, V. (2002). "The mismanagement of customer loyalty". Harvard Business Review, 80(7), 86.

Raorane Abhijit and .Kulkarni R.V, (2011), "DATA MINING TECHNIQUES: A SOURCE FOR CONSUMER BEHAVIOR ANALYSIS", International Journal of Database Management Systems (IJDBMS), Vol. 3, No. 3, pp. 45-56

Sandhya Harikumar , Surya PV ,(2015), "K-Medoid Clustering for Heterogeneous DataSets" , Proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems, Volume 70, Pages 226-237

Venables, W. N. Smith D. M. and the R Core Team,(2016), "An Introduction to R A Programming Environment for Data Analysis and Graphics" , Version 3.3.2 (10-31)

Wei Jainping, (2011), "Research on VIP Customer Classification Rule Base on RFM Model", International Conference on Management Science and Industrial Engineering (MSIE), pp. 336-338