

## بخش بندی بازارهای الکترونیکی بر اساس داده‌های حاصل از رفتار مشتری با استفاده از مدل RFM (مطالعه موردی: شرکت ایران گیت)

عطیه صباغی نیا<sup>a</sup>، بهرام ایزدی<sup>b</sup>

<sup>a</sup> (کارشناس ارشد، گروه مدیریت، دانشکده مدیریت، دانشگاه شیخ بهایی، اصفهان)

<sup>b</sup> (استادیار، گروه مدیریت، دانشکده مدیریت، دانشگاه شیخ بهایی، اصفهان)

نویسنده مسئول: عطیه صباغی نیا

**چکیده:** یکی از راه‌های کسب شناخت در مورد مشتریان، بخش‌بندی آن‌ها به گروه‌های متجانس است. امروزه با پیشرفت تکنولوژی و وجود داده‌های ذخیره شده در پایگاه‌های اطلاعاتی، شرکت‌ها بازاریابی انبوه را کنار گذاشته و به بخش‌بندی بازار از طریق توجه به داده‌های حاصل از رفتار مشتریان روی آورده‌اند. در اینگونه روش‌های بخش‌بندی از داده‌های تراکنشی مشتریان به جای حدس و گمان استفاده می‌شود. این پژوهش داده‌های تراکنشی تازگی خرید، تکرار خرید و ارزش پولی شرکت ایران گیت را با روش‌های داده‌کاوی، خوشه‌بندی کرده و به بررسی رفتار مشتریان بر اساس مدل RFM پرداخته است. نتایج نشان می‌دهد اهمیت پارامتر تازگی خرید و تکرار خرید کاملاً مشهود است و چنین داده‌هایی در جهت تدوین استراتژی‌های مناسب برای هر خوشه راهگشای شرکت‌ها است.

**کلمات کلیدی:** بخش بندی بازار؛ رفتار مشتری؛ مدل RFM.

### ۱. مقدمه

در مسیر تکامل بازاریابی و مصرف و در دنیای کنونی، به دلایلی چون افزایش درآمدها، ارتقا سطح تحصیلات، آگاهی بیشتر و پیشرفت فناوری اطلاعات، شرکت‌ها با نیازها و خواسته‌های متنوع‌تر از سوی مشتریان خود روبرو هستند و با شعار بازاریابی مستقیم روبرو شده‌اند، شعاری که شرکت‌ها را به سمت تمرکز بر مشتری سوق داده است. از این طریق شرکت‌ها قادرند نیازها و ترجیحات مشتریان را شناسایی کرده، آن‌ها را به مشتریان همیشگی خود مبدل نموده و بیشترین رضایت و سوددهی را کسب کنند. بدین منظور در سال‌های اخیر بخش‌بندی بازار با استفاده از تکنیک‌های جدید و همپای علوم دیگر رشد و گسترش یافته که یکی از دستاوردهای مهم آن تعیین صحیح و دقیق بخش‌های بازار، پیش‌بینی ارزش مشتریان با استفاده از پایگاه‌های داده مشتریان و همچنین بررسی و تحلیل رفتار هر یک از مشتریان بر اساس داده‌های ثبت شده در این پایگاه‌ها است. همچنین، تمرکز بر مشتریان فعلی و مطالعه رفتار آنها بهترین نقطه برای شروع یک بازاریابی متناسب با ترجیحات مشتری است.

از این رو شرکت‌های بسیاری بازاریابی انبوه را که به در نظر گرفتن بازار به عنوان یک گروه متجانس می‌پردازد و یک آمیخته بازاریابی برای همه مشتریان پیشنهاد می‌کند، کنار گذاشته و به بازاریابی هدفمند روی آورده‌اند که شامل شناسایی بخش‌های بازار، انتخاب یکی یا چند بخش و تولید محصولات و برنامه‌های بازاریابی منطبق با هر کدام از آن‌ها می‌باشد. هدف‌گیری هوشمندانه به شرکت‌ها کمک می‌کند تا به روشی کارآمد و موثر بر قسمت‌هایی متمرکز شوند که به بهترین شکل می‌توانند خدمت کرده و برایشان نیز بیشترین سودآوری را دارد. این هدف‌گیری به نفع مشتریان نیز هست زیرا شرکت‌ها گروه‌های خاصی از مشتریان را هدف می‌گیرند و پیشنهادهایشان را دقیقاً با نیاز آن‌ها تطبیق می‌دهند [1]. بنابراین، مدیران اهمیت بررسی و تحلیل داده‌های حاصل از رفتار هر یک از مشتریان را با بکارگیری تکنیک‌های بخش‌بندی و در جهت حفظ مشتریان به خوبی درک کرده‌اند [2]. این تقسیم‌بندی بازار به شرکت‌ها کمک خواهد کرد تا با برخورد با ناهمگنی بازار، به ایجاد تعادل میان نیازهای مشتریان متغیر و محدودیت‌های منابع بپردازند و با متمرکز کردن تلاش‌های بازاریابی بر بخش‌های خاص به نحو بهتری نیازهای بازار را پاسخگو باشند. در حقیقت می‌توان گفت هدف اصلی از مطالعه رفتار هر یک از مشتریان ارائه مبنایی برای بخش‌بندی اثر بخش بازار است و بخش عمده‌ای از تحقیقات بازاریابی، مربوط به بخش‌بندی بازارها است. از نقطه نظر مدیران بازاریابی برای توسعه برنامه‌های موفقیت‌آمیز بازاریابی، انتخاب بازار هدف مناسب از اهمیت فوق‌العاده‌ای برخوردار است.

### ۲. ادبیات و پیشینه پژوهش

#### ۱.۲. بخش بندی

بخش‌بندی بازار یک استراتژی مهم برای بازاریابی و شناسایی مصرف‌کنندگان و گروه‌های مختلف با نیازها و علایق گوناگون است. بخش‌بندی بازار کمک می‌کند زیر گروه‌هایی که مصرف‌کنندگان را تشکیل می‌دهند، درک شوند تا بازاریابان بتوانند محصولات و خدمات را بهتر ارائه کنند. می‌توان گفت بخش‌بندی بازار یکی از عوامل اصلی توسعه برنامه‌های بازاریابی موثر می‌باشد و یک روش شناسایی منحصر به فرد برای جدا کردن یک گروه خاص از مصرف‌کنندگان است که رفتارها و نگرش‌های خرید مشابه را دارند. هدف اولیه بخش‌بندی، تقسیم کردن کل بازار به خوشه‌های نسبتاً همگن با الگوهای مصرف مشابه است [3].

در گذشته بسیاری از مدیران به مفهوم بازاریابی انبوه اعتقاد داشتند و بحث بر سر ایجاد بازارهای بزرگ بالقوه بود که به هزینه کمتر و درآمد بیشتر منجر شود. امروزه بسیاری از شرکتها از بازاریابی انبوه به سمت گروههای کوچکتر خریداران با نیازها و ویژگیهای رفتاری مشخص که نیازمند محصولات فردی با آمیخته بازاریابی هستند، حرکت کردهاند [4]. بخشبندی بازار شامل یک طیف گسترده از راهها و روشها می باشد که به دو گروه اصلی تقسیم می شوند: اولین گروه رهیافتهایی هستند که بخشها بر اساس ویژگیهای شناخته شده از یک جمعیت شناخته شده انتخاب می شوند، دومین گروه روشهای Post Hoc هستند که تحقیقاتی تجربی می باشند و با استفاده از تجزیه و تحلیل چند متغیر شناسایی شده در هر بخش در ارتباط هستند [5]. در روشهای قدیمی بخشبندی، خود پژوهشگر تعداد و نوع بخشها را از پیش تعیین و بر مبنای متغیرهایی از پیش تعیین شده مشتریان را به صورت ذهنی گروهبندی می کرد و در نتیجه به دلیل حدس و گمان، ارتباط منسجمی بین بخشهای انتخاب شده توسط پژوهشگر برقرار نبود. اما امروزه بیشتر شرکتها با پیروی از مشتریمداری به بهره گیری از دادهها به روشهای جدید بخشبندی بازار تکیه می کنند. تنها ویژگیهای بخشها را مورد بررسی قرار نمی دهند بلکه به اندازه، رفتار متجانس هر بخش و سودآوری بخشهای بازار در جهت بخشبندی کارا توجه می کنند. در راستای این امر از روشی مناسب استفاده می کنند که فهم آن ساده و از گمانه زنیها به دور باشد.

بخشبندی بازار علاوه بر اینکه شرکتها را به پذیرش واقعیتهای بازار و آمیخته مزایا و منافع زیر را نیز به آنها عرضه می کند:

- فرصتهایی را برای توسعه محصول جدید شناسایی می کند. اغلب اوقات تجزیه و تحلیل دقیقی از بخشهای مختلف مشتریان بالقوه، یک یا چند گروه را آشکار می سازد که نیازهای خاص آنها توسط محصولات فعلی رقبا به خوبی تامین نمی شود. چنین بخشهای پنهان و کشف نشدهای می توانند فرصتهای جذابی را برای طراحی تولید محصولات جدید یا روشهای بازاریابی نوین فراهم آورند.

- به طراحی طرحهای بازاریابی که برای رسیدن به گروههای همگن مصرف کنندگان موثرتر است کمک می کند و برنامههای بازاریابی تفکیکی را به منظور ارضای نیازهای متفاوت خریداران می تواند ایجاد کند.

- فرایند تخصیص استراتژیک منابع بازاریابی را بهبود می بخشد گاهی به منافع استراتژیک بخشبندی توجه کافی نمی شود. اگر محصولاتی خاص برای عرضه در بخشهایی به خوبی مشخص و تعریف شده انتخاب شوند، آنگاه این بخشها به مثابه مراکز سرمایه گذاری یک فعالیت می توانند عمل کنند. بیشتر استراتژیهای تجاری موفق بر مبنای بخشبندی بازار و تمرکز منابع در بخشهای جذابتر استوار می باشند. فرایند بخشبندی باید روی تقسیم بندی بازارها به حوزههایی که در آنها سرمایه گذاری بتواند یک مزیت رقابتی بلند مدت برای شرکت ایجاد کند، متمرکز باشد [6].

## ۲.۲.۲. رفتار مشتری

مطالعه رفتار مشتری معمولاً با شناسایی مشتریان، آیتمهای مکرر رفتار مشتری و الگوهای رفتار خرید آنها سر و کار دارد. این مطالعه، تجزیه و تحلیل مشتریان از جهت اینکه در کجا، چه زمانی و چه چیزی خرید می کنند، می باشد. الگوهای خرید مشتری به نوعی به هم پیوند خورده اند. رفتار خرید مشتریان از طریق عوامل مختلف قابل بررسی است و این عوامل هستند که عادات خرید مشتریان را تحت تاثیر قرار می دهد. مطالعه رفتار مشتری در حل مشکلات بازاریابی و همچنین برای فروش بهتر مفید هستند [7].

بررسی و مطالعه رفتار مشتری، از طریق داده های موجود در پایگاه داده می تواند بینشهای عمیق و جامع از علایق، تجارب و انتظارات مشتریان را فراهم کند. به نوعی تاریخچه فروش یک منبع مهم برای تجزیه و تحلیل رفتار مشتری است. این مطالعه منعکس کننده آیتمهای خریداری شده توسط مشتریان و نمود رفتارهای سر زده از آنها و از دست دادن دیگر رفتارهای مشتریان است [8]. رابطه نزدیکی بین رفتار مشتری و بخشبندی بازار وجود دارد. با درک رفتار مشتری، بخشبندی بازار به درستی انجام خواهد شد. درک رفتار مشتری کار آسانی نیست و نیازمند زمان و بررسی است. موفقیت فعالیتهای بازاریابی تا حد زیادی با برآورده کردن آنچه توسط مصرف کنندگان انتظار می رود تعیین می شود [9].

## ۳.۲. مدل RFM

این مدل یک ابزار قوی و شناخته شده در پایگاههای بازاریابی است که به طور گسترده در اندازه گیری رفتار و ارزش مشتریان بر طبق خرید آنها مورد استفاده قرار می گیرد. مدل RFM اولین بار توسط هوگس در سال ۱۹۹۴ معرفی شد [10]. این مدل یک تکنیک بازاریابی مورد استفاده برای مشتریان مهم مختلف درون پایگاههای داده بزرگ تعیین مشتریان هدف می باشد. تجربه و تحلیل توسط این مدل به بخشبندی مشتریان در پایگاههای داده شرکتها بر اساس رفتار گذشته هر یک از مشتریان کمک می کند. این مدل یک روش شناخته شده برای تحلیل داده های حاصل از رفتار مشتریان و ارزشهای آنها می باشد. با کمی کردن رفتارهای مشتریان، مدل RFM می تواند رفتار مشتری را در پایگاه داده تحلیل و پیش بینی کند. مدل RFM از سه متغیر شامل تازگی خرید (Recency)، تکرار خرید (Frequency) و مبلغ خرید (Monetary) تشکیل شده است. تازگی خرید عبارت است از اندازه گیری تعداد دورهها از زمان آخرین خرید. این شاخص اشاره دارد بر فاصله زمانی بین آخرین خرید صورت گرفته توسط مشتری تا پایان دوره خاص (پایان محدوده زمانی مورد بررسی). کمتر بودن این فاصله نشانگر بالا بودن ارزش این شاخص در مدل می باشد. تکرار خرید عبارت است از، شمارش تعداد خریدهای انجام شده در یک دوره زمانی مشخص که بیشتر بودن تعداد مبادلات، نشانگر بالا بودن ارزش این شاخص در مدل می باشد. مبلغ خرید را می توان به عنوان مبلغ کل پول صرف شده یا مقدار میانگین پول در طول یک دوره زمانی مشخص تعریف کرد که بیشتر بودن مقدار پول صرف شده، بیانگر بالا بودن ارزش این شاخص در مدل می باشد. بنابراین، مدل RFM رفتار خرید مشتری را از طریق بررسی زمانی که مشتریان خرید می کنند، تجزیه و تحلیل می کند [11,12].

### ۴.۲. داده کاوی

با رشد فناوری، شرکت‌ها با مجموعه داده‌های بسیار بزرگ در پایگاه داده روبرو هستند و به دلیل ارزشمند بودن این داده‌ها شرکت‌ها بر آن شده‌اند تا بخش‌بندی و پیش‌بینی که امری غیر قابل اجتناب است را در جهت استفاده مفید از انبوه داده‌های تراکنشی برای شناسایی مشتریان با روش‌های نوین انجام دهند. علیرغم وجود تکنیک‌های گسترده و با وجود پیشرفت امکانات ثبت و ذخیره‌سازی و پردازش داده‌ها، داده‌های خام به تنهایی هیچ کمکی نمی‌کنند و نیاز به روش‌های بهتر، سریع‌تر جهت تحلیل آن‌ها می‌باشد. اگر نتوان مکانیزم‌های کارا و موثری را به منظور استخراج اطلاعات و دانش از داده‌ها طراحی کرد، می‌توان ادعا کرد که سازمان‌ها با فقر دانش روبرو هستند و داده‌ها هم فاقد ارزش خواهند بود. یکی از روشهای موثر در این زمینه داده کاوی است.

داده کاوی که عموماً به عنوان کشف داده در پایگاه داده شناخته می‌شود به استخراج دانش ضمنی مختصر، اطلاعات سودمند بالقوه و ناشناخته گذشته از پایگاه داده اشاره دارد [13]. به نوعی مقدار زیاد داده و عملکرد ناکارای تکنیک‌های سنتی آماری برای داده‌های حجیم، محرکی برای پیدا کردن ابزار بخش‌بندی موثر به منظور کشف اطلاعات مفید در مورد بازارها و مشتریان می‌باشد، بنابراین کشف دانش و داده کاوی یک راه حل برای این مشکل می‌باشد [14]. از این رو، داده کاوی به عنوان یک ابزار قوی به پیدا کردن ارتباط بین قوانین و الگوهای رفتاری از تجزیه و تحلیل مقادیر بزرگ داده اشاره دارد [15]. ایجاد تحول در شناخت داده‌های ذخیره شده در پایگاه های داده و تبدیل آنها به دانش می‌تواند به شرکت ها کمک کند تا تصمیمات بهتری را اتخاذ کنند. داده کاوی تمام عواملی که بر مبنای ارزش مصرف کننده و برآمده از رفتار وی می باشد و در کل دوره مصرف تاثیر می گذارد را در نظر می گیرد. هدف اصلی از استخراج داده‌ها، کشف الگوهای ارزشمند از مجموعه بزرگی از داده‌ها است که مدیران می‌توانند برای تصمیم‌گیری‌های بهتر، بر مبنای رفتار هر یک از مشتریان و همچنین رفتارهای متجانس در هر بخش شناسایی شده، از آن‌ها استفاده کنند [16].

فناوری داده کاوی می‌تواند با قابلیت‌های خود، فرصت جدیدی را برای کسب و کار ایجاد نماید. این قابلیت‌ها عبارتند از:

- پیش‌بینی خودکار روندها و رفتارها: داده کاوی، فرایند یافتن اطلاعات مربوط به آینده را در پایگاه‌های داده بزرگ، خودکار می‌سازد. پرسش‌هایی که در گذشته پاسخگویی به آن‌ها نیازمند تحلیل‌های دستی گسترده ای بود، اینک به سرعت و به صورت مستقیم قابل پاسخگویی می باشد.
- کشف خودکار الگوهای ناشناخته: ابزارهای داده کاوی الگوهایی را که در گذشته پنهان بودند در یک مرحله شناسایی می کنند [17].

### ۳. روش پژوهش

پژوهش حاضر از نظر هدف کاربردی است. از آنجا که این پژوهش داده‌های تراکنش مشتریان را مورد بررسی قرار می‌دهد، جامعه آماری مورد مطالعه، داده‌های مربوط به حدود ۷۰۰۰ نفر از مشتریان شرکت ایران گیت می‌باشد. در پژوهش حاضر به دلیل استفاده از روش‌های داده کاوی از نمونه‌گیری استفاده نشد و تمام داده‌های جامعه مورد استفاده قرار گرفته است و شامل داده های مربوط به سال های ۲۰۰۷ تا ۲۰۱۱ میلادی می باشد. ابزار گردآوری داده ها، داده‌های تراکنشی ذخیره شده مشتریان در پایگاه داده شرکت مربوطه بر مبنای مدل RFM می باشد. از آنجا که در داده کاوی برای انجام محاسبات به یک نرم افزار مناسب نیاز است و امروزه نرم افزارهای مختلفی در این زمینه وجود دارد. در پژوهش حاضر برای انجام محاسبات داده کاوی و اخذ نتایج از نرم افزار Clementine استفاده شده است و سه متغیر تازگی خرید، فراوانی خرید و ارزش پولی برای بخش‌بندی بازار مورد استفاده قرار گرفته است. در پژوهش حاضر برای بررسی روایی نتایج خوشه‌بندی از معیار Silhouette استفاده شده است و به دلیل استفاده از داده‌های تراکنشی ذخیره در پایگاه داده مشتریان شرکت ایران گیت و عدم نیاز به پرسشنامه برای گردآوری داده‌ها، نیازی به بررسی پایایی نبوده است.

بخش ابتدایی شامل گردآوری، آماده‌سازی و پاکسازی داده‌های مزاحم به منظور ارتقا کیفیت داده‌ها است. پس از آن داده‌ها به بازه‌های زمانی یک ساله تفکیک شده‌اند. به دلیل تفاوت در واحد هر یک از شاخص‌ها RFM نرمال‌سازی و یا استانداردسازی داده‌ها انجام شده است. سپس وزن این شاخص‌ها با استفاده از روش تحلیل سلسله مراتبی تعیین شده‌اند. برای تعیین وزن‌های پژوهش حاضر، روش تحلیل سلسله مراتبی گروهی بکار گرفته شده است.

- روش تحلیل سلسله مراتبی: این روش توسط ساعتی، در دهه‌ی ۱۹۷۰ پیشنهاد شد. این روش، مانند آنچه مغز انسان انجام می‌شود به تجزیه و تحلیل مسائل می‌پردازد. تصمیم‌گیرندگان را قادر می‌سازد اثرات متقابل و همزمان بسیاری از وضعیت‌های پیچیده و نامعین را تعیین کنند. این فرایند تصمیم‌گیرندگان را یاری می‌کند تا الویت‌ها را بر اساس اهداف، دانش و تجربه‌ی خود تنظیم نمایند، به نحوی که احساسات و قضاوت‌های خود را به طور کامل در نظر گیرند.

- روش تحلیل سلسله مراتبی گروهی: اگر بخواهیم به جای یک تصمیم‌گیرنده از نظرات دو یا چند تصمیم‌گیرنده استفاده کنیم، در این صورت، از روش AHP گروهی استفاده می‌کنیم. در این‌گونه موارد می‌توان میانگین هندسی نظرات کارشناسان مختلف (X<sup>ij</sup>) را به دست آورد و آن را به عنوان ماتریس اصلی در نظر گرفت. به صورت رابطه (۱) می باشد.

$$X^{ij} = \left( \prod_{l=1}^K X_{ijl} \right)^{\frac{1}{K}} \quad ; \quad i, j = 1, 2, \dots, n \quad i \neq j \quad (1)$$

I شماره تصمیم‌گیرنده، K تعداد تصمیم‌گیرنده و (i, j) شاخص‌ها یا گزینه‌های مورد مقایسه است.  
I=1,2,3,...,K

نکته‌ای که باید در اینجا اشاره کرد این است که اگر بخواهیم ماتریس مقایسات زوجی دارای سازگاری باشد، باید نظرات و مقایسات زوجی که کارشناسان انجام می‌دهند به هم نزدیک باشد. بعد از تشکیل ماتریس مقایسات زوجی، برای شاخص‌ها، بایستی مقادیر آن‌ها را به هنجار کرد. برای این منظور، هر مقدار ماتریس بر جمع ستون مربوطه تقسیم می‌شود. پس از به هنجار کردن، برای محاسبه وزن نسبی هر شاخص میانگین حسابی هر سطر محاسبه می‌شود تا به نتایج نهایی دست یافت. بنابراین هدف از این روش دسترسی به مطمئن‌ترین توافق گروهی خبرگان درباره موضوعی خاص است که با استفاده از نظرخواهی از خبرگان صورت می‌پذیرد [18].

در گام بعد از روش‌های خوشه‌بندی K-means و روش دو مرحله‌ای برای خوشه‌بندی مشتریان استفاده شده است. روش K-means کاربردی‌ترین روش خوشه‌بندی داده‌هاست. این روش اولین بار توسط مک کوپین در سال ۱۹۶۷ ارائه شد. تعداد خوشه‌ها در این روش ثابت و از پیش تعیین شده است. این روش برای خوشه‌بندی داده‌هایی طراحی شد که به صورت عددی (کمی) باشند و خوشه دارای مرکزی به نام میانگین باشد. در این روش، ابتدا اشیا به صورت تصادفی به K خوشه تقسیم می‌شوند. در گام بعد، فاصله هر یک از اشیا از مرکز خوشه خود محاسبه می‌شود. در صورتی که فاصله شی مورد نظر از میانگین خوشه خود زیاد و به خوشه دیگری نزدیک‌تر باشد، این شی به خوشه‌ای که نزدیک‌تر است اختصاص می‌یابد. این کار آنقدر تکرار می‌شود تا تابع خطا حداقل شود و یا اعضای خوشه‌ها تغییر نیابد. اگر D مجموعه داده‌ها با n شی باشد، و C1، C2، ...، Ck بیانگر K خوشه مجزای D باشد در این صورت تابع خطا (EF) مجموع فواصل هر شی از مرکز خوشه خودش تعریف می‌شود. رابطه (۲) در زیر را مشاهده کنید.

$$EF = \sum_{i=1}^k \sum_{X \in C_i} d(X, \mu(C_i)) \quad (2)$$

که در آن  $\mu$  نشان دهنده مرکز (میانگین) خوشه،  $d(X, \mu(C_i))$  فاصله هر شی از مرکز خود است. به دلیل اینکه در خوشه‌بندی‌های مرکز گرا، تابع خطایی وجود دارد که می‌خواهیم آن را حداقل کنیم، می‌توان به مسائل خوشه‌بندی مرکز گرا به دید مسائل بهینه‌سازی نگریست. در این نوع خوشه‌بندی تابع هدفی وجود دارد که تابع خطا است. در خوشه‌بندی K-means، گام‌هایی به قرار زیر طی می‌شود:

گام آغازین: تفکیک داده‌های اولیه به K خوشه به صورت دلخواه

گام تکراری (الف): محاسبه فاصله هر شی از مرکز خود، (ب) محاسبه تابع خطا

گام بهبود: جابجایی عضوی که بیشترین فاصله را با مرکز خوشه خودش دارد، به خوشه‌ای که کمترین فاصله را با آن دارد.

دستور توقف: تغییر نیافتن اعضای خوشه‌ها یا کاهش نیافتن مقدار تابع خطا [19].

در روش دو مرحله‌ای می‌توان هنگامی که هیچگونه اطلاعاتی از چگونگی گروه‌ها وجود ندارد، برای خوشه‌بندی مجموعه داده‌ها به گروه‌های مجزا استفاده کرد. این روش دارای دو مرحله است. مرحله اول با یک گذر از روی داده‌های ورودی، آن‌ها را به زیر خوشه‌های قابل مدیریت فشرده می‌کند. تشخیص داده‌های پرت در این مرحله است. زیر خوشه‌هایی با تعداد رکورد کمتر نسبت به سایر خوشه‌ها به عنوان داده‌های پرت در نظر گرفته می‌شوند و درخت زیرخوشه‌ها با حذف آن‌ها بازسازی می‌شود. برخی از داده‌های پرت محتمل، در صورت شباهت کافی به ویژگی‌های هر یک از زیرخوشه‌های جدید بازسازی شده، می‌توانند به آن‌ها اضافه شوند. بقیه داده‌های پرت محتملی که نمی‌توان آن‌ها را ادغام کرد به عنوان داده‌های پرت شناسایی شده و در یک خوشه مجزا (خوشه پرت) ریخته شده و از مرحله خوشه‌بندی سلسله‌مراتبی حذف می‌شوند. در مرحله دوم از یک روش خوشه‌بندی سلسله‌مراتبی به منظور ادغام تکاملی این زیرخوشه‌ها به خوشه‌های بزرگتر بهره می‌برد. این مرحله نیازی به گذر مجدد از داده‌ها ندارد. خوشه‌بندی سلسله‌مراتبی نیازی به تعیین تعداد خوشه‌ها ندارد و این یک نقطه قوت در این روش است. اکثر روش‌های خوشه‌بندی سلسله‌مراتبی با استفاده از رکوردها به عنوان خوشه‌های اولیه کار خود را آغاز کرده و به طور بازگشتی به منظور تولید خوشه‌های بزرگتر آن‌ها را باهم ادغام می‌کند. به دلیل این‌که چنین روش‌هایی در مواجهه با تعداد زیادی داده شکست می‌خورند، روش پیش خوشه‌بندی دو مرحله‌ای، خوشه‌بندی سلسله‌مراتبی را حتی برای مجموعه‌های بزرگ داده تسهیل می‌کند [20].

اعتبارسنجی خوشه‌بندی، یکی از مسائل بسیار مهم در تحلیل خوشه‌بندی است، به علت اینکه نتایج خوشه‌بندی باید معتبر باشد در بیشتر الگوریتم‌های خوشه‌بندی تعداد خوشه‌ها از عوامل مهم قابل بررسی است که در فرایند پژوهش حاضر از شاخص Silhouette استفاده شده است و بهترین خوشه‌بندی تعیین گردیده است. معیار Silhouette، نسبت متوسط فاصله هر مشاهده به خوشه‌ی دسته خودش به نزدیکترین خوشه یا دسته دیگر را اندازه می‌گیرد. روش اعتبارسنجی Silhouette، براساس میانگین فاصله هر یک از نمونه‌های یک خوشه با تمام نمونه‌های موجود در همان خوشه و میانگین فاصله کل نمونه موجود در خوشه‌های دیگر با یک خوشه مشخص تعریف می‌شود. بر اساس این دیدگاه، برای هر خوشه میزان پراکندگی و همبستگی داده‌ها تعیین می‌شود که مقادیر بیشینه این شاخص برای تعیین تعداد بهینه خوشه به کار می‌رود. معیار Silhouette طبق رابطه (۳) محاسبه می‌شود:

$$S(i) = \frac{(b(i) - a(i))}{\text{Max}\{a(i), b(i)\}} \quad (3)$$

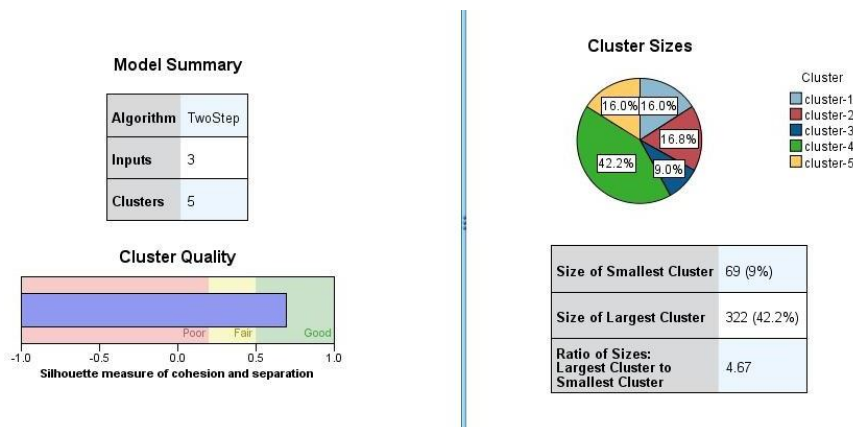
در این رابطه  $b_i$  نمایانگر تشابه نداشتن یک نمونه نسبت به همه نمونه‌های موجود در خوشه‌های دیگر (نزدیکترین خوشه‌ها) است و  $a_i$  نمایانگر تشابه نداشتن یک نمونه با نمونه‌های دیگر در یک خوشه است. مقدار Silhouette بین  $-1$  تا  $+1$  قرار دارد. مقدار مثبت و بزرگ Silhouette یک مشاهده نشان می‌دهد که آن مشاهده به درستی در خوشه مربوط به خود قرار گرفته است. مقدار منفی برای Silhouette یک مشاهده نشان دهنده تخصیص نامناسب آن مشاهده به خوشه است [21]. در ادامه با بررسی تحلیلی هر یک از این سه متغیر رفتاری در مدل RFM استراتژی‌های بازاریابی بر مبنای تحلیل روانشناختی که متناسب بخش‌های شناسایی شده باشد ارائه شده است.

### ۱.۳. یافته‌ها

همانطور که گفته شد، بخش‌بندی پژوهش حاضر بر مبنای داده‌های تراکنشی RFM می‌باشد. پس از آماده‌سازی اولیه سه فیلد تاریخ انجام تراکنش، میزان اعتبار خریداری شده و شماره شناسایی مشتریان مبنای اجرای مدل RFM قرار گرفتند. همچنین با توجه به در اختیار نبودن مبالغ پرداختی مشتریان، فیلد مربوط به اعتبار خریداری شده که معادل مبالغ پرداختی است برای محاسبه ارزش پولی مشتریان مورد استفاده قرار گرفت. برای این منظور سه فیلد مذکور از طریق حذف موارد ناسازگار و پراکنده مورد پردازش قرار گرفتند و بعد از پردازش داده‌های سال‌های ۲۰۰۷ تا ۲۰۱۱ میلادی به کار گرفته شد. در سال‌های ۲۰۰۷ و ۲۰۰۸ میلادی روش دو مرحله‌ای و در سال‌های ۲۰۰۹، ۲۰۱۰، ۲۰۱۱ و روش K-means بالاترین اعتبار و کیفیت را داشته و مبنای محاسبات بعدی قرار گرفته‌اند. نتایج مربوط به بهترین خوشه‌بندی برای هر سال میلادی در جدول‌های ۱ تا ۵ و شکل‌های مربوط به اعتبارسنجی آن‌ها در شکل‌های ۱ تا ۵ آمده است.

جدول ۱. نتایج روش دو مرحله‌ای (سال ۲۰۰۷)

بخش‌ها	تعداد مشتری	R	F	M
۱	۱۲۲	۰,۴۴	۰,۰۰	۰,۰۰
۲	۱۲۸	۰,۸۵	۰,۰۴	۰,۰۰
۳	۶۹	۰,۸۵	۰,۳۷	۰,۲۹
۴	۳۲۲	۰,۷۸	۰,۰۱	۰,۳۳
۵	۱۲۲	۰,۲۲	۰,۰۰	۰,۳۳

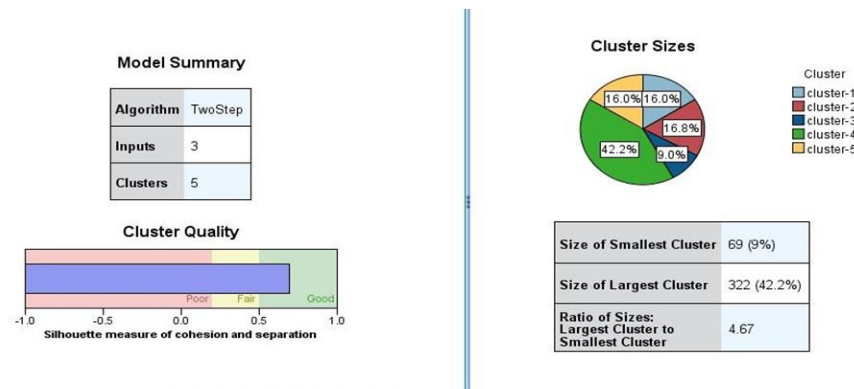


شکل ۱. شاخص Silhouette (سال ۲۰۰۷)

همانطور که در شکل شماره ۱ نشان داده شده در سال ۲۰۰۷ روش دو مرحله‌ای با مقدار Silhouette ۰,۶۹۲ بهترین خوشه‌بندی بوده است. همچنین در این سال خوشه شماره ۳، با M,F,R بالاتر شناسایی شد. همچنین، در سال ۲۰۰۸ روش دو مرحله‌ای با مقدار Silhouette ۰,۶۹۲ بهترین خوشه‌بندی بوده است که در این سال خوشه شماره ۳، با M,F,R بالاتر شناسایی شد.

جدول ۲. نتایج روش دو مرحله‌ای (سال ۲۰۰۸)

M	F	R	تعداد مشتری	بخش‌ها
۰,۰۰	۰,۰۰	۰,۴۴	۱۲۲	۱
۰,۰۰	۰,۰۴	۰,۸۵	۱۲۸	۲
۰,۲۹	۰,۳۷	۰,۸۵	۶۹	۳
۰,۳۳	۰,۰۱	۰,۷۸	۳۲۲	۴
۰,۳۳	۰,۰۰	۰,۲۲	۱۲۲	۵

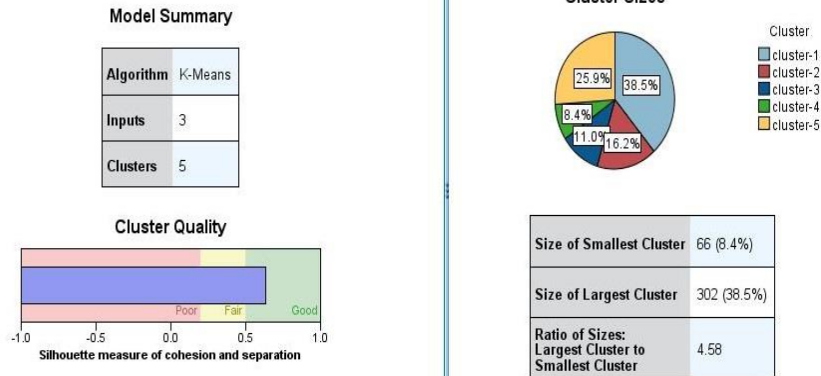


شکل ۲. شاخص Silhouette (سال ۲۰۰۸)

در سال ۲۰۰۹ روش K-means با مقدار Silhouette ۰,۶۳۲ بهترین خوشه‌بندی بوده است که در این سال خوشه شماره ۴، با M.F.R بالاتر شناسایی شد. همچنین، در سال ۲۰۱۰ روش K-means با مقدار Silhouette ۰,۶۳۷ بهترین خوشه‌بندی بوده است که در این سال خوشه شماره ۳، با M.F.R بالاتر شناسایی شد.

جدول ۳. نتایج روش K-means (سال ۲۰۰۹)

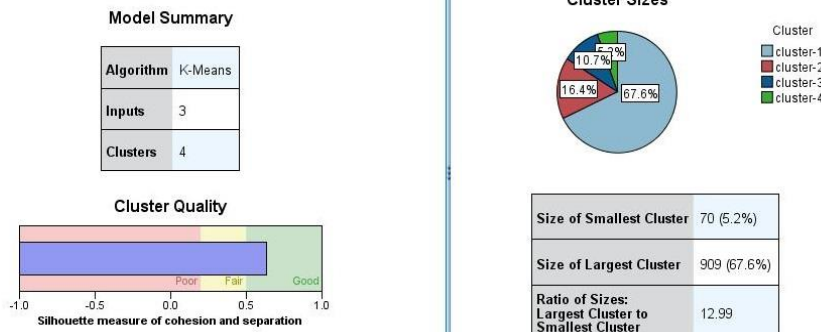
M	F	R	تعداد مشتری	بخش‌ها
۰,۰۲	۰,۰۸	۰,۸۳	۳۰۲	۱
۰,۰۳	۰,۲۹	۰,۴۸	۱۲۷	۲
۰,۴۹	۰,۰۹	۰,۷۹	۸۶	۳
۰,۰۵	۰,۵۲	۰,۸۸	۶۶	۴
۰,۰۲	۰,۰۶	۰,۲۹	۲۰۳	۵



شکل ۳. شاخص Silhouette (سال ۲۰۰۹)

جدول ۴. نتایج روش K-means (سال ۲۰۱۰)

M	F	R	تعداد مشتری	بخش‌ها
۰,۰۱	۰,۰۷	۰,۸۶	۹۰۹	۱
۰,۰۲	۰,۰۶	۰,۳۵	۲۲۱	۲
۰,۰۵	۰,۵۱	۰,۸۹	۱۴۴	۳
۰,۵۱	۰,۰۵	۰,۴۹	۷۰	۴

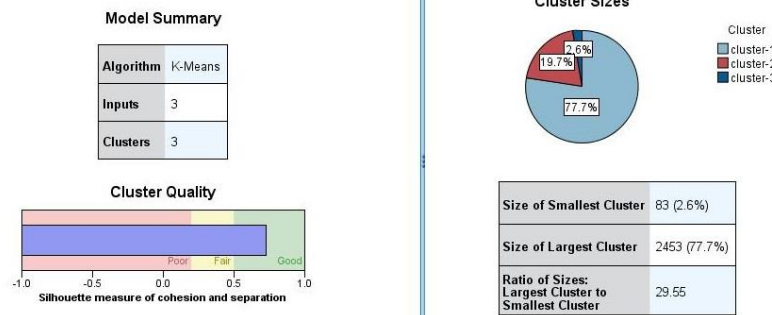


شکل ۴. شاخص Silhouette (سال ۲۰۱۰)

در سال ۲۰۱۱ روش K-means با مقدار Silhouette ۰,۷۲۸ بهترین خوشه‌بندی بوده است که در این سال خوشه شماره ۳، با M.F.R بالاتر شناسایی شد.

جدول ۵. نتایج روش K-means (سال ۲۰۱۱)

M	F	R	تعداد مشتری	بخش‌ها
۰,۰۱	۰,۰۶	۰,۸۶	۲۴۵۳	۱
۰,۰۲	۰,۰۳	۰,۳۶	۶۲۳	۲
۰,۴۹	۰,۰۳	۰,۸۵	۸۲	۳



شکل ۵. شاخص Silhouette (سال ۲۰۱۱)

### ۲.۳. تجزیه و تحلیل یافته‌ها

این پژوهش، برای پاسخ به سوال اینکه چگونه می‌توان مدل RFM را برای بخش‌بندی مشتریان به کار برد؟ و در این راستا و بر اساس مدل RFM چگونه رفتار مشتریان بررسی و تحلیل می‌شود؟ صورت گرفته است. پس از مطالعه و بررسی مشخص گردید که اطلاعات مفیدی از رفتار خرید مشتریان و تغییرات این رفتارها در پایگاه‌های داده ثبت و ذخیره می‌گردد. در صورت هدایت این اطلاعات به سیستم بازاریابی، شرکت‌ها می‌توانند برنامه‌های کارآتری طراحی کنند و از بخش‌بندی سنتی و بر پایه و اساس حدس و گمان فاصله بگیرند. با بکارگیری مدل RFM می‌توان اطلاعات مفیدی از سوابق خرید مشتریان و رفتارهای خرید هر مشتری در طی سابقه خرید او از شرکت استخراج کرد. اطلاعات حاصل از پایگاه داده‌ها و بکارگیری مدل RFM باعث شناسایی مشتریانی که بر اساس نیازها و رفتار خریدشان سودآور هستند، می‌شوند، تا اینگونه شناخت‌ها از مشتریان در راستای ارائه خدمات بهتر و اجرای برنامه‌های بازاریابی موثرتر برای بخش‌های مختلف مشتریان بکارگرفته شود. پس در این مدل مشتریانی که اخیراً خرید کرده‌اند، با تعداد دفعات خرید و مبلغ خرید بالاتر، مشتریان احتمالی تلاش‌های بازاریابی مستقیم شرکت هستند. این مدل در شناخت هر چه بهتر رفتار خرید مشتریان و میزان اهمیت هر یک از پارامترها در صنعت مربوطه و در پژوهش حاضر در بازارهای الکترونیکی راهگشا می‌باشد. در پژوهش حاضر در سالهای مورد بررسی تازگی خرید بالاترین میزان اهمیت را داشته است و بعد از آن تکرار خرید در این بازار الکترونیکی حائز اهمیت است. میزان تازگی خرید و تکرار دفعات خرید مشتریان در این صنعت موجب حفظ رابطه بلند مدت و افزایش وفاداری در این گونه بازارها می‌شود و به همراه آن سودآوری مستمر برای بازار مربوطه حاصل می‌گردد. با توجه به نوع خدماتی بودن این بازار تازگی خرید و همچنین تکرار خرید و حضور مداوم مشتریان باعث دستیابی به اطلاعات بیشتر و استخراج نتایج مفیدتر و با بینش خواهد شد. همچنین با بررسی داده‌ها در دیگر خوشه‌ها، مشتریانی که تازگی خرید پایینی دارند، اگر رویگردان و جذب رقبا نشده باشند، می‌توان با تدوین و اجرای استراتژی‌های پیش‌گیرانه نگهداری نمود. در نهایت می‌توان گفت، هدف از بخش‌بندی مشتریان، ایجاد تمایز بین آن‌ها و تشخیص اینکه مشتریان و رفتارهایشان در چه سطح و اهمیتی قرار دارد و سعی در نگهداری آن‌هاست، زیرا با رشد فناوری و فعالیتهای بازاریابی احتمال رویگردانی مشتریان هم وجود دارد و از آن‌جا که هزینه‌های جذب مشتریان جدید به مراتب بالاتر از هزینه نگهداری مشتریان قدیمی است بکارگیری استراتژی‌های حفظ و اصلاح دید مدیران در این راستا کمک‌کننده است.

### ۴. نتیجه‌گیری

همزمان با رشد چشمگیر تکنولوژی‌های ارتباطی و محاسباتی و سایر دستاوردهای مرتبط و بهره‌گیری روز افزون شرکت‌های کوچک و بزرگ از انبارسازی داده‌ها برای تسهیل و تسریع پردازش، عدم استفاده مدیران از ابزارها و روش‌های مناسب برای درک نیازها و انتظارات مشتریان موجب هدر رفتن منابع و نیز رویگردانی مشتریان از سازمان خواهد شد. از اینرو مزیت رقابتی بادوام در شرکت‌ها از طریق توجه داده‌های ذخیره شده در پایگاه‌ها و بهبوده دانستن آنها، همچنین تحلیل داده‌ها به منظور درک وضعیت و رفتار مشتریان و بخش‌بندی مشتریان در گروه‌های متجانس حاصل می‌شود. یعنی اینکه بتوانند بخش‌های مختلف با هر درجه اهمیت مشخص کنند تا مورد بررسی قرار داده و استراتژی‌های لازم و مناسب را ارائه دهند چرا که شفاف‌سازی اطلاعات هر مشتری و مطالعه رفتار وی برای شرکت‌ها و واحدهای بازاریابی راهگشاست.

### منابع

- [1] کاتلر، ف.، آرمسترانگ، گ. (۱۳۸۹). اصول بازاریابی. ترجمه مهدی زارع. تهران: مدیر فردا.
- [2] Chuang, Y. F., Chia, S. H., & Wong, J. Y. (2013). Customer value assessment of pharmaceutical marketing in Taiwan. *Industrial Management & Data Systems*, 113(9), 1315-1333.
- [3] Shi, T., Liu, X., & Li, J. (2018). Market Segmentation by Travel Motivations under a Transforming Economy: Evidence from the Monte Carlo of the Orient. *Sustainability*, 10(10), 1-11.
- [4] Schejter, A. M., Serenko, A. S., Turel, O., & Zahaf, M. (2010). Policy implications of market segmentation as a determinant of fixed-mobile service substitution: What it means for carriers and policy makers. *Telematics and Informatics*, 27(1), 90-102.



- [5] Hanafizadeh, P., & Mirzazadeh, M. (2011). Visualizing market segmentation using self-organizing maps and Fuzzy Delphi method – ADSL market of a telecommunication company. *Expert Systems with Applications*, 38(1), 198-205.
- [6] واکر، ا. س.، بوید، ه. د.، مالینز، ج.، لرش، ج. ک. (۱۳۹۱). استراتژی بازاریابی. ترجمه محمد اعرابی و داود ایزدی. تهران: دفتر پژوهشهای فرهنگی.
- [7] Rajarageswari, M., & Tamizharasi, T. (2019). Customer Behavior Analysis Using Apriori Algorithm. *Journal of The Gujarat Society*, 21(14), 1210-1214.
- [8] Han, J., Ding, H., Qian, C., Ma, D., Xi, W., Wang, Z., Jiang, Z., & Shangguan, L. (2015). CBID: A Customer Behavior Identification System using Passive Tags. *IEEE/ACM Transactions on Networking*, 24(5), 1-12.
- [9] Rukka, R. M., Amiruddin, A., Tenriawaru, A. N., & Muslimah, C. D. (2020). Effect analysis of market segmentation on purchasing decisions local rice in modern market. *IOP Conf. Series: Earth and Environmental Science*, 575(1), 1-6.
- [10] Han Hua, H., Huang, C. K. T., & Hua Kao., Y. (2013). Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. *The Journal of Systems and Software*, 86, 779-788.
- [11] Wei, T. J., Lin, S. Y., Yang, Y. Z., & Wu, H.H. (2020). Using A Combination of RFM Model and Cluster Analysis to Analyze Customers' Values of A Veterinary Hospital. *IAENG International Journal of Computer Science*, 47(3), 1-7.
- [12] Hu, Y. H., & Yeh, Z. W. (2014). Discovering valuable frequent patterns based on RFM analysis without customer identification information. *Knowledge-Based Systems*, 61(1), 76-88.
- [13] شهرابی، ج.، زارع، ا. (۱۳۹۲). داده‌کاوی با کلمنتاین. تهران: انتشارات جهاد دانشگاهی، واحد صنعتی امیر کبیر.
- [14] Hizioglu, A. (2013). Soft computing applications in customer segmentation: State-of-art review and critique. *Expert Systems with Applications*, 40(16), 6491-6507.
- [15] Cheng, L. C., & Sun, L. M. (2012). Exploring consumer adoption of new services by analyzing the behavior of 3G subscribers: An empirical case study. *Electronic Commerce Research and Applications*, 11(2), 89-100.
- [16] Yoseph, F., & AlMalaily, M. (2019). New Market Segmentation Methods Using Enhanced (RFM), CLV, Modified Regression and Clustering Methods. *International Journal of Computer Science & Information Technology*, 11(1), 43-60.
- [17] توربان، ا.، لیدنر، د.، مک‌لین، ا.، وترب، ج. (۱۳۸۶). فناوری اطلاعات در مدیریت، دگرگونی سازمان‌ها در اقتصاد دیجیتال، (ترجمه حمیدرضا ریاحی، پوریا قطره‌نی، مهدیه توفیقی و حسین صامعی). تهران: چاپخانه دانشگاه پیام نور.
- [18] مومنی، منصور. (۱۳۸۹). مباحث نوین تحقیق عملیات. تهران: منصور مومنی.
- [19] مومنی، منصور. (۱۳۹۰). خوشه‌بندی داده‌ها (تحلیل خوشه‌ای). تهران: مولف.
- [20] علیزاده، س.، ملک‌محمدی، سمیرا. (۱۳۹۲). داده‌کاوی و کشف دانش گام به گام با نرم‌افزار Clementine. تهران: دانشگاه خواجه نصیرالدین طوسی.
- [21] تیمورپور، ب.، سپهری، م. م.، پزشکی، ل. (۱۳۸۸). روشی نوین برای دسته‌بندی هوشمند متون علمی (مطالعه موردی مقالات فناوری نانو متخصصان ایران). فصلنامه علمی پژوهشی، سیاست علم و فناوری، ۲(۲)، ۱-۱۴.