

## بخش بندی مشتریان صنعت دارو براساس مدل RFML

بابک سهرابی<sup>۱</sup>، ایمان رئیسی وانانی<sup>۲</sup>، نسترن نیک آیین<sup>۳</sup>

**چکیده:** در صنعت دارو مدیران بازاریابی و فروش با حجم انبوهی از داده‌های فروش شرکت‌های پخش، به داروخانه‌های مشتری خود مواجه‌اند. یکی از روش‌هایی که به آنان در کنترل وضعیت بازار، رقابت با سایر رقبای، برنامه‌ریزی هر چه بهتر برای افزایش فروش محصولات خود و در نتیجه هدفمند کردن فعالیت‌های بازاریابی کمک خواهد کرد، آگاهی از بخش‌بندی‌های مختلف مشتریان و سیاست‌گذاری بازاریابی و فروش بر مبنای آن خواهد بود. هدف اصلی این مقاله، کمک به مدیران بازاریابی و فروش صنعت دارو، از طریق تعیین و تحلیل بخش‌های مختلف مشتریان و ارائه پیشنهادهای متناسب با هر بخش، به منظور حفظ و افزایش خرید آنان به کمک روش‌های داده‌کاوی است. در این تحقیق، براساس متغیرهای تازگی، تکرار، ارزش پولی و مدت زمان خرید در مدل RFML، داروخانه‌ها در خوشه‌های مختلف قرار گرفته و تحلیل شده‌اند. در نتیجه این بخش‌بندی، سه دسته داروخانه به نام‌های: داروخانه‌های کم‌خرید و کم‌سود، با میزان خرید و سود متوسط و وفادار و پرسود از نظر روند فروش شناسایی شدند و براساس این بخش‌بندی، تحلیل‌های مربوط به آن ارائه شده است.

واژه‌های کلیدی: بخش‌بندی مشتریان، داده‌کاوی، صنعت دارو، مدل RFML

۱. استاد گروه مدیریت فناوری اطلاعات، دانشکده مدیریت دانشگاه تهران، تهران، ایران

۲. استادیار گروه مدیریت، دانشگاه علامه طباطبائی، تهران، ایران

۳. کارشناس ارشد مدیریت فناوری اطلاعات - گرایش هوش تجاری، دانشکده مدیریت دانشگاه تهران، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۵/۰۳/۰۸

تاریخ پذیرش نهایی مقاله: ۱۳۹۵/۰۸/۲۶

نویسنده مسئول مقاله: بابک سهرابی

E-mail: bsohrabi@ut.ac.ir

## مقدمه

امروزه همگام با سایر صنایع، در صنعت دارو<sup>۱</sup> نیز شرکت‌های دارویی به دنبال بهبود اثربخشی محصولات و کاهش مقدار خطا و بازگشت در آن هستند. صنعت داروسازی در سال ۲۰۱۴ تا ۹ درصد و در سال ۲۰۱۵ به میزان ۱۲ درصد رشد کرده است (سنتیل کوماران، مانیکاندان و سنتیل کومار، ۲۰۱۶). هدف غایی مدیران بازاریابی و فروش در هر صنعت، بالابردن سهم فروش آن شرکت در بازار رقابتی است. بخش‌بندی مشتریان<sup>۲</sup> و تحلیل رفتار آنها، در صنایع بزرگ و رقابتی همچون صنعت دارو که به علت تنوع کالاها، بسیار پرگردش هستند، سبب هدفمندشدن فعالیت‌های بازاریابی و ارتباط مؤثر آنها با مشتریان می‌شود (برادران و بیگلری، ۱۳۹۴).

روش‌های تجزیه و تحلیل قدیمی بیش از حد کند هستند، از این رو به اتخاذ تصمیم‌های دقیق مدیریتی منجر نمی‌شوند. مراقبت‌های بهداشتی شامل چالش‌های متنوعی چون تنوع محصولات، نیاز به استفاده ماهرانه از داده‌ها، راه‌حل‌های کسب‌وکار، تصمیم‌گیری‌های پیچیده و انتقال دانش است. در چنین وضعیتی کارکرد سیستم پشتیبانی از تصمیم‌گیری در افزایش بهره‌وری و نظارت بر نتایج، پیدا کردن الگوها، افزایش بهره‌وری فرایند تحویل، حفظ سطح مناسب کیفیت و رسیدن به سود مدنظر، بسیار پررنگ‌تر است (سنتیل کوماران و همکاران، ۲۰۱۶). در این صنعت، مدیران بازاریابی و فروش با حجم انبوهی از داده‌ها مواجه‌اند. این داده‌ها شامل فروش انواع شرکت‌های بخش دارو، به داروخانه‌های مختلف است که به صورت ماهانه به پایگاه داده شرکت‌های دارویی اضافه می‌شوند. از آنجا که تجزیه و تحلیل این داده‌ها کار آسانی نیست و با توجه به فضای رقابتی حاکم بر بازارها، اهمیت پاسخگویی به این نیاز مدیران بازاریابی و فروش، به طور فزاینده‌ای افزایش یافته است. مدیران این صنعت، برای کنترل وضعیت بازار، رقابت با سایر رقبای برنامه‌ریزی برای تأمین موجودی داروخانه‌های سودآور و انبار، تعیین نیاز به ویزیت و مأموریت بیشتر برای نمایندگان علمی، مشخص‌شدن وضعیت مشتریان و در نتیجه تصمیمی مناسب و دقیق، همواره با مشکل مواجه‌اند و برای رفع این نیازها، همواره باید نتیجه تجزیه و تحلیل این داده‌ها را در اختیار داشته باشند و بررسی کنند. یکی از روش‌هایی که در این موقعیت به آنان کمک خواهد کرد، اطلاع از بخش‌های مختلف مشتریان است. بخش‌بندی مشتریان از رویکردهای داده‌کاوی<sup>۳</sup> است که به کشف گروه‌های مشابه از مشتریان می‌انجامد؛

- 
1. Pharmaceutical Industry
  2. Customer Segmentation
  3. Data mining

بیشتر براساس متغیرهای تازگی<sup>۱</sup>، تکرار<sup>۲</sup> و ارزش پولی خرید<sup>۳</sup> در مدل RFM انجام می‌شود (برادران و بیگلری، ۱۳۹۴). چنانچه از بخش‌بندی مناسبی استفاده نشود، برای همه مشتریان شرایط برابری قرار داده می‌شود که می‌تواند مشکلاتی را در این زمینه ایجاد کند.

اهمیت تحقیق از بعد نظری این است که با افزایش سرعت تحول در علوم، ضرورت استفاده از دانش‌های نوین بیش از پیش محرز شده است. داده‌کاوی به‌عنوان رشته علمی نوین در زمینه بازاریابی و استخراج اطلاعات می‌تواند، تأثیر شایان توجهی در دستیابی به این اهداف داشته باشد. الگوریتم‌های داده‌کاوی بسیار متنوع و پرکاربرد هستند، اما مقالات پژوهشی و مقالات مفهومی مربوط به استفاده از داده‌کاوی در صنعت دارو و به‌ویژه استفاده از آن در حیطه بازاریابی و فروش، بسیار اندک است و همین مسئله انگیزه‌ای برای شروع تحقیق حاضر بوده است. در این مقاله با استفاده از بخش‌بندی مشتریان و تحلیل ویژگی‌های هر بخش، اطلاعاتی فراهم خواهد شد که می‌تواند مبنای تصمیم‌گیری مدیران برای برنامه‌ریزی و تعیین هدف برای آینده باشد. ضرورت و اهمیت تحقیق از جنبه کاربردی نیز در این است که با استفاده از نتایج این تحقیق، اطلاعات و شاخص‌های لازم برای تصمیم‌گیری و برنامه‌ریزی مدیران بازاریابی و فروش سازمان‌های دارویی، میان حجم عظیمی از داده‌های موجود، فراهم خواهد شد. به کمک این بخش‌بندی و توصیه‌های ارائه‌شده برای رفتار با هر بخش، یک الگوی علمی و عملی در اختیار مدیران بازاریابی و فروش سازمان‌های دارویی قرار می‌گیرد که می‌تواند آنها را در رسیدن به پاسخ سؤالات و ارائه پیشنهاد‌های مناسب و کاربردی یاری رساند و در نتیجه به ارتقای تصمیمات هوشمندانه مدیریتی و بهبود مزیت رقابتی در صنعت دارو منجر شود. به‌طور کلی، می‌توان دلیل اجرای پژوهش حاضر را این‌گونه خلاصه کرد، کمک به مدیران بازاریابی و فروش صنعت دارو، از طریق تعیین و تحلیل بخش‌های مختلف مشتریان و ارائه پیشنهاد‌های متناسب با هر بخش، برای حفظ و افزایش خرید آنان به کمک روش‌های داده‌کاوی.

### پیشینه پژوهش

با توسعه فناوری اطلاعات، یافتن اطلاعات مفید از مجموعه وسیع داده‌ها، به موضوع مهمی تبدیل شده است. برای این منظور داده‌کاوی مشهورترین تکنیک است که به‌طور موفقی در

1. Recency
2. Frequency
3. Monetary

بسیاری از حوزه‌ها به‌مثابه ابزار تحلیل به‌کار گرفته شده است. داده‌کاوی اطلاعات ناواضح و از پیش ناشناخته و بالقوه را از داده‌ها استخراج می‌کند (چی یو، یی فنگ، تینگ و هه چون، ۲۰۰۹).

### پیشینه نظری پژوهش

در ادامه، بر مبنای نکات یادشده در بخش‌های قبل، به مفاهیم مربوط به بخش‌بندی مشتریان، داده‌کاوی و کمکی که به تصمیم‌گیری مدیریتی در صنعت دارو می‌کند و همچنین مدل RFML پرداخته می‌شود.

### بخش‌بندی مشتریان

بخش‌بندی بازار به این منظور انجام می‌شود که برای هر یک از بخش‌های بازار، راهبردهای مناسبی توسعه داده شود. بخش‌بندی بازار، فرایندی است که بازار را به زیربخش‌های متمایزی از مشتریانی که نیازها و مشخصاتشان یکسان است، دسته‌بندی می‌کند. تقسیم بازار سبب می‌شود که یک بازار نامتجانس<sup>۱</sup> به خرده‌بازارهای متجانس گروه‌بندی شود. بخش‌بندی، یکی از مهم‌ترین مقوله‌ها برای رسیدن به بازاریابی مدرن و مدیریت موفق ارتباط با مشتری است و از آن به‌عنوان نخستین گام در طراحی راهبرد بازاریابی مشتری‌محور یاد می‌شود (دهدشتی و پورحسینی، ۱۳۹۲).

### داده‌کاوی

داده‌کاوی فرایند شناسایی الگوهای معتبر، جدید، مفید و قابل فهم در بین داده‌هاست. فرایند داده‌کاوی به‌کمک کار روی داده و ساده‌کردن تحلیل آغاز می‌شود. اولین و پذیرفته‌ترین تعریف داده‌کاوی تبدیل داده‌های خام به اطلاعات است. داده‌کاوی از داخل داده‌های خام، ابرداده<sup>۲</sup> و دانش خاص مشتری را استخراج می‌کند (آرورا، دیپالی و شیپرا، ۲۰۱۵). در تعریفی دیگر، وظیفه کشف الگوها را در مجموعه داده‌های بزرگ برعهده دارد که شامل روش‌های هوش مصنوعی، یادگیری ماشینی، آمار و پایگاه داده می‌شود. داده‌ها ممکن است شامل داده‌های مکانی، داده‌های چندرسانه‌ای، داده‌های سری زمانی، داده متنی و داده‌های وب باشند. داده‌کاوی فرایند استخراج الگوهای جالب، کوچک اما بااهمیت، ضمنی، ناشناخته و بالقوه مفید یا دانش از مقدار بسیار عظیم داده‌هاست. این فرایند دربردارنده مجموعه‌ای از فعالیت‌ها برای پیدا کردن الگوهای جدید، ضمنی یا غیرمنتظره در داده‌ها یا الگوهای غیرمعمول در آنهاست. داده‌کاوی می‌تواند به سازمان‌ها به‌منظور

1. Heterogeneous market  
2. Metadata

درک بهتر کسب‌وکار، ارائه خدمات بهتر به مشتریان و افزایش اثربخشی سازمان در بلندمدت کمک کند (مهتا و گاوانده، ۲۰۱۵). رایج‌ترین توابع مدل‌سازی در الگوریتم‌های داده‌کاوی موجود عبارت‌اند از طبقه‌بندی<sup>۱</sup>، رگرسیون، خوشه‌بندی، قوانین انجمنی<sup>۲</sup>، ایجاد قانون، خلاصه‌سازی، مدل‌سازی وابستگی (تحلیل لینک) و تحلیل دنباله (کشف توالی) (چن و چن، ۲۰۰۷). تکنیک‌های خوشه‌بندی از راه‌شناسایی گروه‌هایی از کاربران که به‌نظر می‌رسد اولویت‌های مشابهی داشته باشند و دسته‌بندی گروه‌هایی که اولویت‌های بسیار متفاوتی دارند، کار می‌کنند. به‌طور خلاصه، روش خوشه‌بندی خوب، خوشه‌هایی را با کیفیت زیاد تولید می‌کند، به‌طوری که درون هر کلاس بیشترین شباهت و بین کلاس‌های متفاوت کمترین شباهت وجود داشته باشد. با این حال، خوب بودن یک خوشه، درنهایت به نظر کاربر وابسته است. خوشه‌بندی اقدامی است که یک جمعیت همگن را به چند خوشه ناهمگن دسته‌بندی می‌کند (خاتمی فیروزآبادی و صادقیانی، ۲۰۱۵). خوشه‌بندی روش بدون ناظر نامیده می‌شود. این روش معمولاً برای آن دسته از بخش‌ها و خوشه‌هایی استفاده می‌شود که در تجزیه و تحلیل‌های آینده به‌کار گرفته خواهند شد، مانند دسته‌بندی مشتریان به بخش‌های مشابه که هر بخش ویژگی‌ها و طرح بازاریابی منحصر به فرد خود را دارد (نگای، هو، وونگ، چن و سان، ۲۰۱۱). الگوریتم k-medoids از روش‌های بخش‌بندی شناخته‌شده برای خوشه‌بندی داده‌هاست. در این روش گروه‌های داده‌ای بر مبنای نزدیکی به یکدیگر بر اساس فاصله اقلیدسی انتخاب می‌شوند (آرورا و همکاران، ۲۰۱۵). این الگوریتم مبتنی بر شیء است و نماینده خوشه‌ها را از میان خود داده‌ها و نه میانگین‌گیری از آنها انتخاب می‌کند. در واقع medoid یا میانه یک خوشه، مرکزی‌ترین عنصر خوشه است. هدف این روش کم‌کردن حساسیت نسبت به مقادیر بزرگ در مجموعه داده‌هاست. الگوریتم PAM مشابه k-means با k عدد تصادفی از میان داده‌ها شروع می‌شود و در هر مرحله در صورتی که مجموع فاصله‌ها درون خوشه بهبود یابد، نزدیک‌ترین داده به مقدار میانگین به‌عنوان مرکز خوشه جدید انتخاب می‌شود. این روش برخلاف k-means در برابر داده‌های پرت و نویزدار مقاوم‌تر است؛ ضمن آنکه در این الگوریتم، ابتدا ماتریس فاصله‌ها (فاصله داده‌ها از یکدیگر) محاسبه می‌شود، پس سربار<sup>۳</sup> محاسباتی برای محاسبه فاصله بردار داده‌ها از مرکز خوشه‌ها وجود ندارد (ولمورگان و سنتانم، ۲۰۱۰؛ پرادپ و سینگ، ۲۰۱۱).

طبقه‌بندی از رایج‌ترین روش‌های کاربردی داده‌کاوی است که مجموعه‌ای از نمونه‌های از پیش طبقه‌بندی‌شده را برای توسعه مدلی که قادر به طبقه‌بندی جامعه بزرگ باشد، به‌کار می‌برد.

1. Classification
2. Association Rules
3. Overload

برنامه‌های کاربردی تشخیص تقلب و مخاطرات اعتباری، به‌خوبی با این نوع از تجزیه و تحلیل تناسب دارند (مهتا و گاوآنده، ۲۰۱۵). الگوریتم درخت طبقه‌بندی یادگیری، نوعی ساختار درختی است که همه گره‌های غیر برگ با یک ویژگی و هر گره برگ با یک طبقه برچسب‌گذاری شده است. کار از گره ریشه آغاز می‌شود؛ در هر گره، داده‌ها به صورت پی‌درپی و با شناسایی ویژگی‌ها و آستانه تقسیم، بخش‌بندی می‌شوند و طبقات بهتر در زیرمجموعه‌های مربوط به داده‌ها جای می‌گیرند (جیانگینگ، ینگجیان و یانگوانگ، ۲۰۰۹).

### داده‌کاوی در صنعت دارو

در صنعت داروسازی، تکنیک‌های داده‌کاوی در یکی از مدل‌های پروفایل‌سازی، طبقه‌بندی، خوشه‌بندی، پیش‌بینی و تجزیه و تحلیل داده‌ها به‌کار می‌رود. تجزیه و تحلیل داده‌ها در حوزه دارو، به‌کمک استنتاج از تحقیقات داده‌کاوی به تعیین عملکرد بهینه تصمیمات کسب‌وکارها در آینده کمک می‌کند. امروزه بیشترین استفاده از علم داده‌کاوی در صنعت دارو در حوزه پیش‌بینی براساس تجویز گذشته پزشکان است. بسیاری از مؤسسه‌های بهداشتی فاقد سیستم‌های اطلاعات مناسب برای تولید گزارش‌های قابل اعتماد با توجه به اطلاعات موجود هستند، اما داده‌کاوی می‌تواند در رسیدن به این گزارش‌ها بستر لازم را مهیا کند (پرینس و استگوی، ۲۰۰۰). در جوامع اطلاعاتی پسا صنعتی، داده‌ها یکی از منابع باارزش محسوب می‌شوند. این داده‌ها مفید نیستند، مگر اینکه پردازش شده و به اطلاعات تبدیل شوند و مزیت رقابتی ایجاد کنند. صنایع دارویی کارآمد نیز به‌طور فزاینده داده‌هایی در فرمت‌های مختلف تولید می‌کنند. در این موقعیت و میان کوهی از داده‌ها، چگونه می‌توانیم فرصت‌ها و روندها را شناسایی کنیم و به استنتاج نتایج پردازیم؟ در این میان، علم داده‌کاوی می‌تواند داده‌ها را به اطلاعات تبدیل کرده و به پیمودن موفق این مسیر کمک کند. تکنیک‌های داده‌کاوی به سازمان‌ها اجازه می‌دهد تا این پدیده‌ها را بشناسند و با مشکل سربار اطلاعات به‌خوبی برخورد کنند. تکنیک‌های داده‌کاوی برای ایجاد ارزش افزوده در صنایع دارویی، به‌خوبی توانسته‌اند خود را با این کسب‌وکارها همسو کرده و بازگشت سرمایه را برای این صنایع بیمه کنند (رانجان، ۲۰۰۹).

### فروش دارو

در گذشته جریان داده‌های صنعت دارو کمتر و مدیریت آن ساده‌تر بود؛ اما امروزه به سمت پیچیده‌تر شدن حرکت می‌کند و استفاده از تکنولوژی‌های اطلاعاتی در آن ضروری است

محصولات دارویی اغلب محصولات ضروری برای مراقبت و سلامتی هستند؛ اما بسیاری از محصولات دارویی در بازار جایگزین دارند. از این رو، مدیران شرکت‌های دارویی همواره باید به فکر استراتژی‌های جدید برای تسخیر بازار دارو باشند. نمایندگان علمی در واقع همانند فروشندگان دارو عمل می‌کنند و در نتیجه تیم فروش دارو، باید واجد شرایط لازم باشند و اغلب از داروسازان، پزشکان و محققان بازاریابی تشکیل (سنتیل کوماران و همکاران، ۲۰۱۶).

### مدل RFML

یکی از مدل‌های مطرح در بخش‌بندی و تحلیل ارزش مشتری، مدل بخش‌بندی موسوم به RFM<sup>۱</sup> است که هاگ در سال ۱۹۹۴ آن را ارائه کرده است. در این مدل، سه متغیر تازگی خرید (R)، تکرار خرید (F) و ارزش مادی خرید (M) به‌عنوان متغیرهای عملکردی مشتریان برای بخش‌بندی آنها پیشنهاد شده است (وی، لینگ و ونگ، ۲۰۱۲). مدل RFM معیارهای مؤثری برای تجزیه و تحلیل رفتار مصرف‌کنندگان و بخش‌بندی آنها فراهم می‌کند. در مدل یادشده سه متغیر برای تعیین کمیت مشتری، وفاداری و سهم آن استفاده می‌شود. بر پایه آن مشتریان بخش‌بندی می‌شوند و اطلاعات بسیار مفیدی به‌منظور تصمیم‌گیری به‌دست می‌آید (چانگ و پی تسای، ۲۰۱۱). امروزه مدل RFM توسعه پیدا کرده و پارامتر طول ارتباط با مشتری (Length) هم به آن اضافه شده است که این پارامتر به فاصله بین اولین و آخرین خرید اشاره می‌کند (زارع حسینی و محمدزاده، ۲۰۱۶).

### پیشینه تجربی پژوهش درباره بخش‌بندی مشتریان بر اساس مدل RFML

هو و یه در سال ۲۰۱۴ مدلی به نام RFM-customer-patterns برای کشف رفتارهای مشتری در صنایع خرده‌فروشی ارائه کردند. آنها با ترکیب مدل RFM و الگوهای مورد علاقه خرید، رفتار مشتریان را بررسی کردند. لی، دانگ و تسین (۲۰۱۱) در تحقیق خود پارامتر (L) طول مدت ارتباط را به مدل RFM اضافه کردند و نشان دادند به‌کمک این پارامتر دقت بخش‌بندی افزایش می‌یابد و بهتر می‌شود. سوینی و فتحعلی‌زاده در سال ۲۰۱۲ از مدل RFM برای بخش‌بندی مشتریان بیمه استفاده کردند. آنها در این فرایند مدل RFM را به LRFMC توسعه دادند که این کار موجب افزایش کیفیت بخش‌بندی برای مشتریان بیمه شد. در این مقاله L نشان‌دهنده طول مدت همکاری و C هزینه مشتری است. سیدحسینی، ملکی و غلامیان در سال ۱۳۹۰ پژوهشی با عنوان «تحلیل خوشه‌بندی با استفاده از رویکرد داده‌کاوی به‌منظور توسعه روش‌شناسی

1. Recency, Frequency, Monetary

CRM در ارزیابی وفاداری مشتریان» اجرا کردند. این تحقیق روش RFM که یکی از روش‌های تحلیل میزان وفاداری و ارزش مشتریان است را با وزن‌دهی کردن توسعه می‌دهد. چانگ و پی تسای در سال ۲۰۱۱ در تحقیقی مدل RFM را به‌عنوان مدلی معرفی کردند که دارای معیارهای مؤثری برای تجزیه و تحلیل رفتار مصرف‌کنندگان و بخش‌بندی آنها فراهم می‌کند. آنها به‌کمک این مدل و سه متغیر آن، به تعیین کمیت وفاداری مشتری و سهم آن پرداختند و اطلاعات بسیار مفیدی به‌منظور تصمیم‌گیری فراهم کردند. زارع حسینی و محمدزاده در سال ۲۰۱۶ در مقاله خود به کشف دانش در رفتار بیماران به کمک داده‌کاوی پرداختند، آنها مدل RFM را توسعه دادند و به آن پارامتر طول ارتباط با مشتری (Length) را اضافه کردند. این پارامتر در تحقیق یادشده به فاصله بین اولین و آخرین ویزیت اشاره می‌کند.

### پیشینه تجربی پژوهش درباره بخش‌بندی مشتریان

حسین‌زاده، کرمی و مهربانی در مقاله‌ای با عنوان «بخش‌بندی مشتریان در رستوران‌های زنجیره‌ای براساس سبک تغذیه» مشتریان رستوران‌های زنجیره‌ای فست‌فود بوف را در شهر تهران بخش‌بندی کردند. آنها در این پژوهش به‌کمک روش تحلیل عاملی اکتشافی و تحلیل خوشه‌ای، به تجزیه و تحلیل نتایج پرداختند. در نتایج این تحقیق، سه گروه مشتریان با عناوین آگاه از سلامت، بی‌علاقه‌ها و راحت‌طلبان، براساس سبک تغذیه در رستوران‌های زنجیره‌ای شناسایی شدند (حسین‌زاده، کرمی و مهربانی، ۱۳۹۳). برادران و فرخی (۱۳۹۳) در مقاله‌ای، بر پایه مدل توسعه‌یافته RFM و با افزودن متغیر توالی روزهای انجام تراکنش (C)، الگوی جدیدی برای بخش‌بندی مشتریان ارائه دادند. در این مدل مشتریان بانک براساس مدل RFM و مدل پیشنهادشده (RFMC) و با استفاده از الگوریتم خوشه‌بندی دومرحله‌ای و به‌کارگیری گام‌های متدلوژی GRISP-DM بخش‌بندی شده‌اند. آنها در پژوهش خود نشان دادند دقت مدل RFMC نسبت به مدل RFM در بخش‌بندی مشتریان این صنعت به مقدار ۵/۵ درصد بیشتر است. در این راه از الگوریتم شبکه عصبی استفاده کردند. (ملکی و حاجیلو، ۱۳۹۵) در مقاله‌ای مشتریان کلیدی بخش اینترنت مخابرات استان قم را بخش‌بندی کردند. در این پژوهش پس از تعیین مقادیر RFM (تازگی مبادله، تعداد دفعات مبادله و ارزش پولی مبادله) و وزن هر یک از این شاخص‌ها براساس فرایند تحلیل سلسله‌مراتبی، مشتریان به روش دومرحله‌ای خوشه‌بندی شدند. نتایج به‌دست‌آمده زمینه را برای تحلیل ویژگی‌های مشتریان و اولویت‌بندی آنان فراهم کرد. اسفیدانی، محمدی، کیماسی و پارسافر (۱۳۹۳) در مقاله‌ای به بخش‌بندی بازار بانکداری خرد برمبنای مزایای مورد انتظار مشتریان از طریق مطالعات میدانی و مصاحبه اکتشافی پرداختند. در نتیجه تحلیل خوشه‌ای چهار بخش مشخص شد که از میان آنها، بخش سوم بزرگ‌ترین بخش بود. این



بخش برای عواملی چون کارکنان و اعتماد، راحتی، سهولت خدمات و تکنولوژی، اهمیت بسیاری قائل است. آنها در نهایت متناسب با ویژگی‌های رفتاری هر بخش، راهکاری برای بازاریابی ارائه کردند. صالح‌آبادی (۲۰۱۵) در پژوهشی با عنوان «دسته‌بندی مشتریان صنعت پخش دارو به کمک روش‌های داده‌کاوی»، با استفاده از داده‌کاوی و تکنیک شبکه‌های عصبی به پیش‌بینی رفتار مشتریان آتی در یکی از شرکت‌های پخش دارو پرداخت. در نتیجه این پژوهش مشتریان به سه بخش تفکیک شدند و ویژگی‌های هر یک از بخش‌ها بررسی و تجزیه و تحلیل شد. در پایان نیز پیشنهادهایی برای هر یک از گروه‌های مشتریان مطرح شده است. زارع حسینی و محمدزاده (۲۰۱۶) به کمک تکنیک‌های داده‌کاوی بیماران یک بیمارستان را بخش‌بندی کردند. آنها ابتدا با افزودن پارامتر طول، مدل RFM را به RFML توسعه دادند، سپس با بهره‌مندی از الگوریتم‌های K-means، Two-step و tree (CHAID Decision)، بخش‌های مختلف بیماران را شناسایی کردند. در نتیجه دانش به‌دست‌آمده از این بخش‌بندی توانستند CRM بیمارستان مد نظر را تقویت کنند. کادروس و دومینگوئس (۲۰۱۴) در پژوهشی با هدف شناسایی ارزش واقعی هر مشتری برای تخصیص منابع به صورت هدفمند، چارچوبی را برای بخش‌بندی مشتریان ارائه دادند و در آن ارزش دوره عمر مشتری، ارزش فعلی و وفاداری مشتری را در نظر گرفتند؛ سپس بخش‌های مشتریان را به کمک نقشه خودسازمانده‌ی شده تعیین کردند. آنها این مدل را روی مشتریان یک آسیاب نیشکر پیاده‌سازی کردند و در نتیجه آن نه بخش به‌منظور تصمیم‌گیری شناسایی شد.

پاداشی (۱۳۹۲) در پژوهشی به بخش‌بندی مشتریان یکی از شرکت‌های سرویس‌دهنده خدمات اینترنتی پرداخت. هدف او داده‌کاوی مشتریان برای انجام بخش‌بندی مشتری و در نتیجه تعریف استراتژی‌های مناسب و مفید برای پیروزی در محیط رقابتی برای این شرکت‌ها بود. او در این پژوهش برای بخش‌بندی از مدل RFM استفاده کرد و پس از تعیین مقادیر شاخص‌های مدل RFM شامل تازگی مبادله، تعداد دفعات مبادله و ارزش پولی مبادله در میان مشتریان شرکت یادشده، روش‌هایی برای بخش‌بندی مشتری و همچنین توانایی آنها در شناسایی رفتارهای مشتریان، به‌ویژه ریسک لغو خدمات شرکت ارائه داد.

کلاروزکی، تنگلر و مجرکاکوا (۲۰۱۶) مشتریان را در زمینه خدمات پستی بخش‌بندی کردند. آنان برای طراحی نوعی نرم‌افزار مدیریت ارتباط با مشتری (CRM) مناسب برای سازمان‌های پستی، به‌دنبال رویکرد نوینی برای بخش‌بندی مشتریان بودند. از نظر محققان بخش‌بندی درست مشتریان می‌تواند بخشی از زنجیره ارزش CRM باشد. در این مقاله ضمن توصیف روش پایه‌ای

برای بخش‌بندی مشتریان، یک مدل پیشرفته CRM براساس ماتریس چندبعدی طراحی شده است.

سسیکوا، آنتوسوا و میهالکوا (۲۰۱۵) به‌منظور بهبود امکانات و ارائه خدمات در بازار در حال رشد و محیط رقابتی حمل‌ونقل هوایی، مشتریان فرودگاه سریلانکا را بخش‌بندی کردند. مهم‌ترین هدف این تحقیق، توجه بیشتر به بخش‌های حمل‌ونقل هوایی به‌منظور بهبود موقعیت بازار، رشد اقتصادی و جذب بیشتر سرمایه‌گذاران بود. آنها با در نظر گرفتن فرضیه‌هایی، رابطه خدمات ارائه‌شده و رضایت مشتریان را از طریق اطلاعات جمع‌آوری‌شده از پرسشنامه بررسی کردند. در نتیجه این تحقیق، مشتریان در گروه‌های مختلف دسته‌بندی شدند و پیشنهادهایی برای ارائه خدمات بهتر به هر بخش مطرح شد.

پاتاک، لستاکوا، کوردوا و ولکوا (۲۰۱۴) مشتریان داروخانه‌های الکترونیکی را با توجه به اطلاعات جمع‌آوری‌شده از فروشگاه‌های الکترونیکی بخش‌بندی کردند. آنها معتقدند با توجه به افزایش کاربران الکترونیکی و محیط رقابتی کسب‌وکارهای الکترونیکی، برای حفظ مشتریان باید به رویکردهای نوین روی آورد. در نتیجه این تحقیق، بخش‌های مختلف و عواملی برای تشویق مشتریان به استفاده از فروشگاه‌های الکترونیکی شناسایی و معرفی شده است.

### پیشینه تجربی پژوهش درباره استفاده از داده‌کاوی در صنعت دارو

سنتیل کوماران و همکارانش (۲۰۱۶)، در مقاله‌ای تأثیر داده‌کاوی در صنعت دارو و انواع مختلف تکنیک‌های آن را که می‌تواند در صنعت دارو استفاده شود، بررسی کردند. در این مقاله درباره یافتن الگوهای تکراری از طریق روابط انجمنی، طبقه‌بندی و پیش‌بینی داروها در آینده، تکنیک‌های خوشه‌بندی برای دارو و بازار دارو، اهمیت داده‌کاوی و آمار و ثبت آزمایش صحبت شده است. رانجان در سال ۲۰۰۹ به مزایای داده‌کاوی در صنعت دارو و کمک به بخش‌های مختلف آن پرداخت و با مطالعه مروری تحقیقات، مشاهدات، گزارش‌ها و نکات قبلی، مزایای استفاده از تکنیک‌های داده‌کاوی را در صنعت دارو بررسی کرد. با توجه به تحقیقات وی، این تکنیک‌ها در پیش‌بینی فروش دارو و کشف داروی جدید بیشترین کاربرد را دارند. در طول این مقاله به مزایای تکنیک‌های شبکه عصبی، دسته‌بندی، خوشه‌بندی و روابط انجمنی در صنعت دارو پرداخته شده است.

لاروس در سال ۲۰۰۹، در مقاله‌ای با موضوع کشف دانش در داده به کمک داده‌کاوی، به این موضوع پرداخته است که استفاده از تکنیک‌های مختلف داده‌کاوی می‌تواند مشکلات کسب‌وکارها را برطرف کند و سبب رسیدن به مزیت رقابتی شود. سادارینا، کوتاری و گوندالیا (۲۰۱۳) تکنیک‌های داده‌کاوی برای بازاریابی محصولات دارویی را پیاده‌سازی کردند. در طول این

تحقیق اطلاعات از منابع مختلف جمع‌آوری شده و به کمک الگوریتم‌های درخت تصمیم و روش طبقه‌بندی، تجزیه و تحلیل شده و در نهایت به گرفتن تصمیمات مهم در بازاریابی دارویی کمک کرده است. صالح‌آبادی در سال ۱۳۹۴ تکنیک شبکه‌های عصبی را در یکی از شرکت‌های پخش دارو به کار برد. هدف این مقاله استفاده از روش شبکه‌های عصبی به منظور دسته‌بندی مشتریان صنعت پخش دارو و در نهایت ارزیابی نتایج به دست آمده است. طبق نتایج به دست آمده، مشتریان به سه بخش تفکیک شدند و ویژگی‌های هر یک از بخش‌ها بررسی و تجزیه و تحلیل شد.

هیسترز در سال ۲۰۰۹ در تحقیقات خود به این موضوع می‌پردازد که داده‌کاوی در صنعت دارو تبدیل به نوعی کسب‌وکار شده است. در این کسب‌وکار اطلاعات مرتبط به صنعت دارو جمع‌آوری و به همه شرکت‌های نیازمند این اطلاعات فروخته می‌شود. شرکت‌های دارویی اطلاعاتی را خریداری می‌کنند که به آنها در هدف‌گذاری بهتر برای فروش کمک خواهد کرد. امروزه استفاده از داده‌کاوی در این صنعت در حال افزایش است زیرا پتانسیل بسیار زیادی از خود نشان داده است و به شرکت‌های دارویی در افزایش راندمان کاری و سودآوری کمک می‌کند.

رانجان در سال ۲۰۰۷ به این موضوع پرداخت که فهم دقیق دانش نهفته در اطلاعات شرکت‌های دارویی، برای حفظ جایگاه رقابتی و تصمیم‌گیری‌ها حیاتی است. شرکت‌های دارویی اغلب بدون سیستم‌های اطلاعاتی مناسب هستند و سیستم‌های موجود فقط برای نگهداری اطلاعات استفاده می‌شوند و نمی‌توانند بسیاری از گزارش‌های مورد نیاز مدیران را تولید کنند. حجم اطلاعات دارویی زیاد است و این موضوع از وجود محصولات دارویی متنوع در بازار پیداست؛ اما سیستمی وجود ندارد که این اطلاعات را به دانش تبدیل کند. هدف این مقاله استخراج الگوهای مفید از میان داده‌های انبوه و بهبود کیفیت تصمیم‌گیری در کشف دارو، پیش‌بینی، آزمایش و توزیع آن است.

رای و شوپا (۲۰۱۰) در مقاله‌ای انواع مختلف تکنیک‌های خوشه‌بندی را بررسی کردند. آنها ابتدا انواع خوشه‌ها را تشریح کردند و سپس شاخص‌هایی را به منظور تحلیل تکنیک‌های خوشه‌بندی ارائه دادند. در نهایت، ضعف‌ها و قوت‌های آنها را تعیین کردند که در این راه از الگوریتم‌های سلسله‌مراتبی، تفکیک (افراز)، چگالی‌محور و شبکه‌های عصبی استفاده کردند.

سید حسینی و همکارانش (۲۰۱۰) مقاله‌ای با عنوان «تحلیل خوشه‌بندی با استفاده از رویکرد داده‌کاوی به منظور توسعه روش‌شناسی CRM در ارزیابی وفاداری مشتریان» را ارائه کردند. این تحقیق ابتدا روش RFM که یکی از روش‌های تحلیل میزان وفاداری و ارزش مشتریان است را با وزن‌دهی کردن (WRFM) توسعه می‌دهد. سپس داده‌های مرتبط با مشتریان را جمع‌آوری می‌کند و با استفاده از تکنیک K-means که یکی از روش‌های داده‌کاوی است، به خوشه‌بندی مشتریان اقدام می‌کند. در مرحله بعد ارزش مشتریان را از هر دو روش RFM و WRFM

محاسبه کرده و نتایج را با هم مقایسه می‌کند. نتیجه این تحقیق نشان داده است که ادغام دو روش WRFM و K-means به بهبود تعیین میزان وفاداری مشتریان منجر شده است.

فانگا ونگ در سال ۲۰۱۱ مقاله‌ای با عنوان «سیستم پیش‌بینی فروش در صنعت پخش تنباکوی چین» ارائه کرد. او در این تحقیق سه دپارتمان فروش، برنامه‌ریزی و خرید را پیش‌بینی کرد. بدین ترتیب که واحد فروش از مدل پیش‌بینی خاکستری، واحد برنامه‌ریزی از مدل چرخه عمر محصول و واحد خرید از مدل میانگین متحرک و ضرایب فصلی استفاده می‌کنند، سپس به هر یک از این واحدها وزنی اختصاص داده می‌شود. در نهایت، پیش‌بینی فروش کل تنباکو با ترکیب سه روش یادشده به دست می‌آید.

جانسون در سال ۲۰۱۰ کتابی با عنوان پیش‌بینی فروش در صنعت دارو نوشت که می‌توان گفت بهترین و کامل‌ترین تحقیقی که در زمینه پیش‌بینی فروش در صنعت دارو صورت گرفته، مربوط به این کتاب است. در این کتاب اصول شش‌گانه یک پیش‌بینی خوب ارائه شده و سپس نتایج تحقیقاتی که در حوزه پیش‌بینی در صنعت دارو انجام گرفته بررسی شده و هر یک از مدل‌ها و روش‌های پیش‌بینی در صنعت دارو را براساس اصول شش‌گانه تجزیه و تحلیل کرده است. این روش‌ها شامل پیش‌بینی شهودی، مدل‌های برون‌یابی، مدل‌های بخش‌بندی، مدل‌های خودگردان‌سازی، مدل‌های اقتصادسنجی و مدل‌های علت و معلولی می‌شود.

بهبودی، مینایی و ابراهیم‌پور (۱۳۹۱) پژوهشی با موضوع «ارائه یک روش داده‌کاوی ترکیبی برای سگمنت‌بندی مشتریان براساس رفتارهای خرید» ارائه دادند. متداول‌ترین الگوریتم خوشه‌بندی که برای مجموعه داده‌های مشتریان به کار می‌رود الگوریتم K-means است؛ اما مشکل اصلی این است که الگوریتم حساسیت آن به انتخاب مراکز اولیه هر خوشه و احتمال همگرایی به راه حل بهینه محلی است. در این مقاله برای کاهش این مشکل و بهبود سگمنت‌بندی مشتریان براساس رفتارهای خرید آنها از ترکیب دو الگوریتم PSO و K-means استفاده و عملکرد الگوریتم ترکیبی با استفاده از مجموعه داده‌هایی از فروشگاه زنجیره‌های مواد غذایی ماندربین ارزیابی و خوشه‌های به دست آمده از این روش از دیدگاه رفتارهای خرید مشتریان تحلیل شده است نتایج ارزیابی بیان می‌کند که در روش پیشنهادی میزان خطای خوشه‌بندی به گونه چشمگیری کاسته شده است.

### روش‌شناسی پژوهش

این پژوهش از نظر هدف کاربردی، به لحاظ روش به دلیل بررسی وضعیت موجود و استفاده از داده‌های حقیقی و توصیف آنها توصیفی و از منظر نوع داده‌ها کمی است. جامعه هدف این پژوهش داده‌های داروخانه‌های مشتری یک شرکت دارویی در داخل ایران است. در این تحقیق

به منظور انجام عملیات داده‌کاوی و مدل‌سازی داده‌ها از روش‌شناسی CRISP-DM<sup>۱</sup> (چاپمن و همکاران، ۲۰۰۰) استفاده شده است که شامل گام‌های زیر می‌شود (شکل ۱):



شکل ۱. گام‌های پژوهش

**شناخت کسب‌وکار:** در این مرحله به بررسی و شناخت کسب‌وکار شرکت‌های تولیدکننده و واردکننده دارو، فرایند فروش دارو در ایران و این شرکت‌ها پرداختیم تا با اهداف، عوامل موفقیت و دغدغه‌های مدیران بازاریابی و فروش این کسب‌وکار آشنا شویم.

**شناخت داده‌ها:** در این مرحله داده‌های فروش شرکت‌های دارویی مد نظر را شناختیم. این داده‌ها به صورت ماهانه در یک سیستم اطلاعاتی مانند برنامه مدیریت اطلاعات بازاریابی (Maris) وارد و ذخیره می‌شوند. در این پایگاه داده به‌ازای هر پخش طرف قرارداد، فروش به صورت ماه به ماه منتقل می‌شود. در هر فایل به‌ازای هر داروخانه‌ای که در آن ماه خرید داشته است، داده‌های فروش آن قابل مشاهده‌اند.

**آماده‌سازی داده‌ها:** در این گام داده‌های مربوط به فروش یک سال (۱۳۹۳-۱۳۹۴) یکی از شرکت‌های دارویی بزرگ و فعال در ایران را انتخاب کردیم. تعداد کل داده‌ها ۱۳۸۲۰ رکود بوده و به‌ازای هر رکود، این ویژگی‌ها تعریف شده بودند: کد پخش، نام پخش، کد مرکز پخش، نام مرکز پخش، کد داروخانه، نام داروخانه، کد رتبه، نام رتبه، کد استان، نام استان، کد شهر، نام شهر، کد منطقه جغرافیایی، نام منطقه جغرافیایی، مجموع مرجوعی، متوسط قیمت خرید، نام محصولات، اولین تاریخ خرید، آخرین تاریخ خرید، تعداد خرید در آخرین خرید، مبلغ خرید در آخرین خرید، متوسط ارزش پولی مجموع خرید. در مراحل پیش‌پردازش داده در میان این داده‌ها پنج رکود بودند که مشخصه «نام شهر»، «استان» و «منطقه» برای آنها بدون مقدار بود که این پنج رکود را پاک کردیم. با توجه به تعداد زیاد ویژگی‌ها و افزایش پیچیدگی مدل، از میان ویژگی‌های موجود در این دیتاست، چهار ویژگی RFML «تازگی خرید»، «تکرار خرید»، «ارزش پولی خرید» و «طول مدت ارتباط» در نرم‌افزار Microsoft Sql Server تولید شدند. متغیر «تازگی خرید» برابر تعداد روزهای بین آخرین خرید تا ۱۳۹۴/۱۲/۲۹، «تعداد تکرار خرید» برابر تعداد کل دفعات خرید در یک سال، «ارزش پولی خرید» برابر مجموع ارزش ریالی خرید در یک سال و متغیر «طول مدت ارتباط» برابر تعداد روزهای بین اولین و آخرین خرید است.

**مدل‌سازی:** برای شروع مدل‌سازی فیله‌های این مجموعه داده در نرم‌افزار رییدماینر وارد شد. این داده‌ها ابتدا وارد الگوریتم‌های خوشه‌بندی و سپس پیش‌بینی شدند. در بخش خوشه‌بندی مشتریان، به کمک تکنیک‌های خوشه‌بندی  $k$ -means،  $k$ -medois، DBSCAN و X-means داده‌ها خوشه‌بندی و بهترین مدل انتخاب شد.

### یافته‌های پژوهش

پس از انجام خوشه‌بندی و تحلیل خوشه‌ها برپایه دسته‌بندی‌های انجام‌گرفته، در این بخش به تشریح ویژگی‌های خوشه‌ها براساس یافته‌های تحقیق پرداخته می‌شود. تحلیل خوشه‌هایی که در ادامه بیان خواهند شد و پیشنهادها متناسب با هر بخش می‌توانند در شناسایی بخش‌های مختلف داروخانه‌ها و تدوین استراتژی متناسب با رفتار هر بخش، به مدیران بازاریابی و فروش شرکت‌های دارویی کمک کنند. نتایج این بخش‌بندی در جدول ۱ قابل مشاهده است. بعد از خوشه‌بندی با این الگوریتم‌ها مشخص شد که درباره این داده‌ها عملکرد دو الگوریتم  $K$ -means و X-means کاملاً شبیه هم بود و نتایج یکسانی به‌دست آمد. الگوریتم DBSCAN نتوانست این مجموعه داده را درست خوشه‌بندی کند و برای این کار مناسب نبود. در نهایت، هر سه الگوریتم  $K$ -means، X-means و  $K$ -medois سه خوشه را مناسب تشخیص دادند.

جدول ۱. نتایج خوش‌بندی و ارزیابی مدل‌ها

نام الگوریتم	تعداد خوشه	اعضای خوشه	زمان اجرا	دیویس-بولدین	سیلوئت
K-Means	۳	خوشه ۱	۳ ثانیه	۰/۷۸۹	۰/۲۱۶
		خوشه ۲			
		خوشه ۳			
K-Medois	۳	خوشه ۱	۶۰ دقیقه	۰/۹۷۳	۰/۳۳۰
		خوشه ۲			
		خوشه ۳			
X-Means	۳	خوشه ۱	۳ ثانیه	۰/۷۸۹	۰/۲۱۶
		خوشه ۲			
		خوشه ۳			
DBSCAN	۲	خوشه ۱ خوشه ۲	۲ دقیقه	...	۱

**ارزیابی و اعتبارسنجی مدل:** پس از مدل‌سازی نتایج به دست آمده از مدل ارزیابی می‌شود. با استفاده از ابزار ارزیابی خوشه‌ها در رپیدماینر، مدل‌های خوشه‌بندی، K-medois، X-means و K-means را با دو معیار Silhouette و Davies Bouldin ارزیابی کردیم. برای اعتبارسنجی مدل خوشه‌بندی DBSCAN نیز، از اپراتور Cluster Distance Performanc استفاده کردیم و در نهایت نتایج طبق جدول ۱ به دست آمد. در جدول ۱، مقدار شاخص دیویس-بولدین و سیلوئت به ازای هر یک از الگوریتم‌ها محاسبه شده‌اند و کمتر بودن شاخص دیویس-بولدین و بالاتر بودن عدد شاخص سیلوئت، برای مدل K-medois، نشان‌دهنده کیفیت بهتر این خوشه‌بندی نسبت به سایر الگوریتم‌هاست. همچنین برای اطمینان از انتخاب بهترین k برای الگوریتم K-medois، اعداد (۷، ۵، ۳ و ۹) را برای این الگوریتم به کار گرفتیم و ارزیابی کردیم، که بهترین k، عدد ۳ انتخاب شد و مطابق نتایج به دست آمده در جدول ۲، هر چه تعداد خوشه‌ها زیادتر می‌شد کیفیت خوشه‌های ایجاد شده کمتر می‌شد.

به منظور حفظ و به کارگیری دانشی که بعد از خوشه‌بندی ایجاد شده است، می‌توانیم از روش‌های طبقه‌بندی استفاده کنیم. در این مرحله با بهره‌گیری از الگوریتم‌های طبقه‌بندی به پیش‌بینی خوشه داروخانه‌های جدید پرداختیم. قبل از اعمال الگوریتم‌های طبقه‌بندی، ابتدا داده‌ها را با نسبت ۳۰-۷۰ به مجموعه آموزش و تست تقسیم کردیم و سپس به پیش‌بینی داده‌ها توسط الگوریتم‌های پیش‌بینی پرداختیم. نتایج به دست آمده از اجرای الگوریتم‌های پیش‌بینی برای مشتریان جدید، به شرح جدول ۳ است.

جدول ۲. انتخاب بهترین K برای الگوریتم K-Medois

K	دیویس - بولدین	سیلوئت	زمان اجرا
۳	۲/۷۹۵	۰/۲۶۶	۵۰ دقیقه
۵	۲/۱۸۷	۰/۱۱۶	۲:۲۰ دقیقه
۷	۲/۱۳۸	۰/۰۶۷	۳:۲۵ دقیقه
۹	۲/۱۰۴	۰/۰۲۳	۱۰:۳۱ دقیقه

جدول ۳. نتایج به دست آمده از اجرای الگوریتم‌های پیش‌بینی برای مشتریان جدید

r2	Error(RMSE)	Accuracy	Precision	Recall	Algorithm
۰/۱۶۴	۰/۴۰۵	%۸۱/۲۵	%۶۲/۷۱	%۷۳/۹۸	Decision Tree
۰/۱۶۲	۰/۴۰۲	%۸۱/۲۸	%۷۱/۰۷	%۶۷	Decision Tree (ID3)
۰/۱۶۲	۰/۴۰۳	%۸۲/۸۲	%۷۳/۷۵	%۷۴/۰۲	Naïve Bayes
۰/۱۳۹	۰/۳۷۳	%۸۴/۴۱	%۷۵/۶۶	%۷۶/۳۳	Neural Net(Auto MLP)
۰/۱۶۵	۰/۴۰۶	%۸۱/۱۹	%۶۶/۹۳	%۶۸/۷۲	Decision Tree (Random Forest)
۰/۱۳۸	۰/۳۷۲	%۸۴/۴۸	%۷۵/۹۳	%۷۴/۴۱	Neural Net
۰/۲۷۷	۰/۵۲۷	%۷۲/۲۶	%۵۶/۴۹	%۵۴/۳۵	KNN

نتایج این جدول نشان می‌دهند تقریباً همه الگوریتم‌ها برای پیش‌بینی مشتریان جدید خوب عمل می‌کنند و صحت پیش‌بینی زیادی را ارائه می‌دهند، که از میان آنان برای پیش‌بینی خوشه مشتریان جدید، الگوریتم Neural Net عملکرد بهتری نسبت به سایرین دارد و خطایی آن کمتر است، به طوری که ۸۴/۴۸ درصد صحت پیش‌بینی را ارائه می‌دهد.

**مرحله توسعه:** در این فاز به کمک مدل به دست آمده، گزارشی به نام RFML به گزارش‌های برنامه سیستم مدیریت اطلاعات بازاریابی (Maris) اضافه شد. تا مدیران علاوه بر سایر گزارش‌هایی که در نتیجه وجود داده‌های فروش از این سیستم دریافت می‌کردند، از این پس بتوانند بخش‌های مشتریان خود را ببینند و تحلیل کنند.



### نام‌گذاری و تحلیل خوشه‌ها

در این قسمت با توجه به رفتار خرید مشتریان و برای سهولت در دسته‌بندی داروخانه‌های جدید، خوشه‌ها را نام‌گذاری می‌کنیم. داروخانه‌های درون هر خوشه براساس شاخص‌های تازگی، تکرار، ارزش پولی و مدت زمان خرید در مدل RFML به همراه نرخ مرجوعی محصولات، در سه دسته داروخانه‌های کم‌خرید و کم‌سود، با میزان خرید و سود متوسط و وفادار و پرسود قرار گرفتند. در ادامه، به توصیف و تحلیل ویژگی‌های داروخانه‌های درون هر خوشه می‌پردازیم.

**خوشه داروخانه‌های کم‌خرید و کم‌سود:** این خوشه داروخانه‌هایی هستند که خریدشان کمتر از سایر داروخانه‌هاست و همچنین میزان مرجوعی بیشتری نیز دارند. بیشترین اعضای این خوشه داروخانه‌های متعلق به استان خراسان رضوی و شهر مشهد هستند. این داروخانه‌ها اغلب رتبه C دارند. نسبت به داروخانه‌های دو خوشه دیگر، از نظر مدت زمانی که مشتری ما هستند در رده دوم قرار می‌گیرند. نسبت به داروخانه‌های دو خوشه دیگر مدت‌زمان طولانی‌تری است که از ما خرید نمی‌کنند. متوسط تعداد تکرار خرید این داروخانه‌ها کمتر از سایر داروخانه‌هاست. در کسب‌وکار دارو معمولاً از این داروخانه‌ها با نام «سرد» یاد می‌کنند.

**خوشه داروخانه‌های با خرید و سود متوسط:** این خوشه بزرگ‌ترین خوشه است و اغلب داروخانه‌ها در این بخش قرار می‌گیرند. این داده‌ها داروخانه‌هایی هستند که از نظر متوسط خرید و نرخ مرجوعی در رده دوم قرار می‌گیرند. بیشترین اعضای این خوشه داروخانه‌های متعلق به استان تهران و شهر تهران هستند. این داروخانه‌ها اغلب رتبه C دارند. از نظر طول مدت زمانی که مشتری ما هستند، بین دو خوشه دیگرند. داروخانه‌های این بخش از نظر تازگی خرید و همچنین تکرار خرید در رده دوم قرار می‌گیرند.

**خوشه داروخانه‌های وفادار و پرسود:** این خوشه کوچک‌ترین خوشه است. مشتریان در این بخش، داروخانه‌هایی هستند که متوسط خریدشان از ما خیلی بیشتر از سایر داروخانه‌ها و متوسط نرخ مرجوعی محصولات نیز خیلی کمتر است. این داروخانه‌ها جزء مشتریان طولانی‌مدت ما هستند و از آخرین خریدشان (از نظر تازگی خرید) از ما به‌طور متوسط فقط ۳۰ روز می‌گذرد. این مشتریان در واقع پرسودترین مشتریان ما هستند به‌طوری که متوسط خریدشان از محصولات ما بسیار بیشتر از سایر مشتریان است. بیشترین اعضای این خوشه اغلب رتبه A دارند و داروخانه‌های متعلق به استان آذربایجان شرقی و شهر تبریز هستند. در کسب‌وکار دارو معمولاً از این داروخانه‌ها با نام «برتر» یاد می‌کنند.

## نتیجه‌گیری و پیشنهادها

در این پژوهش داروخانه‌های مشتری شرکت دارویی مد نظر، بخش‌بندی شده و خوشه‌های آن شناسایی شدند. سپس با طبقه‌بندی امکان پیش‌بینی بخش مشتریان جدید به‌وجود آمد. این سیستم آمادگی بخش‌بندی مشتریان فعلی و پیش‌بینی مشتریان جدید را برای کمک به مدیران بازاریابی و فروش، به منظور اتخاذ تصمیم، برنامه‌ریزی برای طرح‌های بازاریابی و فروش و شناسایی مشکلات بالقوه را دارد. همچنین این سیستم در همه شرکت‌های بازاریابی دارویی دیگر و براساس داده‌های فروش آنها قابلیت پیاده‌سازی و تحلیل نتایج را دارد. همان طور که در بخش پیشینه تجربی بررسی شد، تا کنون مطالعات متعددی با موضوع بخش‌بندی مشتریان صورت گرفته است و این مطالعات در صنایع مختلف انجام شده‌اند؛ اما در این میان سهم صنعت رقابتی و پویای دارو و به‌ویژه فروش آن، بسیار ناچیز است. در این زمینه و در صنعت دارو، در ایران فقط یک مطالعه با موضوع «دسته‌بندی مشتریان صنعت پخش دارو به کمک روش‌های داده‌کاوی» صورت گرفته است که مشتریان پخش دارو را بخش‌بندی می‌کند. در اغلب مطالعات پیشین برای بخش‌بندی، پارامترهای مدل RFM شامل تازگی مبادله، تعداد دفعات مبادله و ارزش پولی مبادله مبنا قرار داده شده است؛ اما در این پژوهش با الهام‌گرفتن از نتایج پژوهش‌های قبلی، این مدل با افزودن پارامتر طول ارتباط با مشتری (Length) به RFML تغییر یافته و در تحلیل در نظر گرفته شده است. علاوه بر این، در پژوهش حاضر ضمن مشخص کردن بخش‌های مشتریان، به منظور کمک به تصمیم‌گیری مدیران بازاریابی و فروش این حوزه برای برنامه‌ریزی و تخصیص منابع به‌صورت بهینه، پیشنهادهای کاربردی متناسب برای هر بخش از مشتریان ارائه شده است.

## پیشنادهایی برای خوشه داروخانه‌های کم‌خرید و کم‌سود

اغلب این داروخانه‌ها مقیم شهرستان هستند و خرید کمتری دارند و از محصولات ما استقبال چندانی نمی‌کنند. ابتدا می‌توان به تحلیل رقبا پرداخت، تا علت تمایل به محصولات آنان و ترجیح داده‌شدن به محصولات ما مشخص شود. بعد از تحلیل رقبا و یافتن دلایل نفوذ آنان در این بخش از بازار، مدیران بازاریابی و فروش می‌توانند متناسب با دلایل این روی‌گردانی و استقبال نکردن، استراتژی مناسبی را در پیش بگیرند و برنامه‌های بازاریابی و فروش هوشمندانه‌تری را تدوین کنند. به‌طور مثال، مدیر بازاریابی می‌تواند برای این داروخانه‌ها استراتژی ویزیت چهره‌به‌چهره<sup>۱</sup> را برنامه‌ریزی کند تا از مهارت بازاریابان خود برای ترغیب این داروخانه‌ها به سمت

1. Face to Face

محصولات خود کمک بگیرد. البته، با توجه به اینکه اغلب این داروخانه‌ها مقیم شهرستان هستند، روش یادشده کمی هزینه‌بر خواهد بود.

### **پیشنادهایی برای خوشه داروخانه‌های با میزان خرید و سود متوسط**

این بخش از مشتریان، بزرگ‌ترین بخش هستند. در گام اول می‌توان برای کاهش هزینه و تحت پوشش قراردادن بخش بیشتری از این داروخانه‌ها از روش ویزیت تلفنی کمک گرفت و در صورت مثبت‌بودن نتیجه برای برخی از این مشتریان، گام بعدی می‌تواند استراتژی ویزیت چهره‌به‌چهره باشد. مدیران می‌توانند برای این بخش از مشتریان یک طرح بازاریابی<sup>۱</sup> در نظر بگیرند. به‌طور مثال، پیشنهاد بسته ریالی می‌تواند انگیزه بدهد. همچنین برای پیشی‌گرفتن از رقبا می‌توان در محل این داروخانه‌ها، ایستگاه‌های تبلیغاتی با حضور نمایندگان فروش خود احداث کرد.

### **پیشنادهایی برای خوشه داروخانه‌های وفادار و پرسود**

این داروخانه‌ها از مشتریان قدیمی ما هستند و از محصولات ما راضی‌اند. اولین گام برای مشتریان این خوشه تلاش برای حفظ آنهاست. این مشتریان ارزش پولی بسیاری را برای ما ایجاد می‌کنند. برای حفظ این داروخانه‌ها بهتر است از نمایندگان فروش ارشد<sup>۲</sup> برای ویزیت مشتریان کلیدی استفاده کنیم. استراتژی مناسب برای آنان، ایجاد وفاداری مطلوب در مشتری، یا همان بازاریابی رابطه‌مند نام دارد که عبارت‌است از همه اقداماتی که شرکت‌ها برای آگاهی و ارائه خدمت بهتر به تک‌تک مشتریان با ارزش خود انجام می‌دهند. این امر غلبه رقبا را با ترندهایی همچون قیمت‌های فروش کمتر و تلاش برای روی‌گردانی مشتری، دشوار می‌سازد (کاتلر، ۱۳۸۴: ۸۵). در گام بعدی می‌توان برای آنان هدف‌گذاری از نوع سید خرید تعریف کنیم تا به خرید محصولات بیشتری از ما تشویق شوند. همچنین می‌توان در محل این داروخانه‌ها غرفه‌های اعطای نمایندگی یا امتیاز<sup>۳</sup> برپا کرد.

الگوریتم‌های داده‌کاوی بسیار متنوع و پرکاربرد هستند، اما مقالات پژوهشی و مقالات مفهومی مربوط به استفاده از داده‌کاوی در صنعت داروسازی بسیار نادر است و بیشتر پژوهش‌های قبلی به‌منظور استفاده از این علم برای کشف داروی جدید است. با توجه به قابلیت‌های این علم و تنوع جریان داده‌های صنعت دارو، مدیران بازاریابی و فروش می‌توانند به

---

1. Merchandising Plan  
2. Key Account Manager  
3. Franchising

کمک آن به انواع تحلیل‌ها و اطلاعات کاربردی در این زمینه دست یابند. به‌عنوان پیشنهاد برای پژوهش‌های آینده می‌توان به کمک الگوریتم‌های داده‌کاوی، به بررسی تأثیر ویژگی‌هایی که برای خود داروخانه‌ها و پزشکان اطراف این داروخانه‌ها صورت می‌گیرد روی فروش داروخانه‌ها پرداخت و همچنین رابطه بین عملکرد نمایندگان علمی<sup>۱</sup> و تأثیر آن بر فروش داروخانه‌ها را بررسی کرد.

## References

- Allahyari Soeini, R. & Fathalizadeh, E. (2012). Customer segmentation based on modified RFM model in insurance industry. *International Journal of Modeling and Optimization*, 25, 101-104.
- Arora, P., Deepali, D. & Shipra, V. (2015). Analysis of K-Means and K-Medoids Algorithm for Big Data. *Procedia Computer Science*, 78, 507-512.
- Asosheha, A., Bagherpour, S. & Yahyapour, N. (2008). Extended acceptance models for recommender system adaption, case of retail and banking service in Iran. *WSEAS transactions on business and economics*, 5(5), 189-200. (in Persian)
- Baradaran, V. Farokhi, Z. (2014). "Customer Segmentation in the Banking Industry by Extended Model of RFMC. *Journal of Brand Management*, 1(2), 135-154. (in Persian)
- Baradaran, V. & Biglari, B. (2015). Customer Segmentation in the Distribution Industry. *Journal of Brand Management*, 1(7), 23-42. (in Persian)
- Behboudi, M., Minae, B. & Ebrahimpour, H. (2012). A data mining method for customer segmentation According to Buy behavior. *International Conference on Management and Economics Industrian engineering. Mahshahr*. (in Persian)
- Blocker, C. & Flint, D. J. (2007). Customer segments as moving targets: integrating customer value dynamism into segment instability logic. *Industrial Marketing Management*, 36(6), 810– 822.
- Chang, H-C. & Tsai, H-P. (2011). Group RFM analysis as a novel framework to discover better customer consumption behavior. *Expert Systems with Applications*, 38(12), 14499–14513.

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. & Wirth, R. (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. USA.
- Chen, C. & Chen, A. (2007). Using data mining technology to provide a recommendation service in the digital library. *The Electronic Library*, 25(6), 711-724.
- Chiu, C. Y., Chen, Y.F., Kuo, I.T. & Ku, H.C. (2009). An intelligent market segmentation system using k-means and particle swarm optimization. *Expert Systems with Applications*, 36(3), 4558-4565.
- Csikosova, A., Antosova, A. M. & Mihalcova, B. (2015). Segmentation of Airports' Customers in Slovakia. *Procedia Economics and Finance*, (23), 1068-1083.
- Cuadros, A. D. & Domínguez, V. (2014). Customer segmentation model based on value generation for marketing strategies formulation. *Estudios Gerenciales* (30), 25-30.
- Dehdashti, SH. & Pourhossein, A. (2012). Performance Implications of Sales & Marketing Strategy. *Quarterly Journal of Business Management*, (5), 61-84. (in Persian)
- Dianjun, F. & Weibing, W. (2011). Sales Forecasting System for Chinese Tobacco Wholesalers. *Procedia Environmental Sciences*, (11), 180 – 186.
- Esfidani, R. M., Mohamadi, M., Keimasi, H., Parsafar, M. (2014). Retail banking market segmentation based on the expected benefits of Bank Mellat customers. *Quarterly Journal of Business Management*, 2(6), 227-250. (in Persian)
- Heesters, D. M. (2009). An assault on the business of pharmaceutical data mining. *journal of pennsylvania journal of business law*, 11(3), 789-821.
- Hosseinzade, SH., Karami, M. & Mehrabani, M. (2013). Customers segmentation in chain restaurants based on nutrition style. *Quarterly Journal of Business Management*, 7, 83-99. (in Persian)
- Hu, Y. & Yeh, W. (2014). Discovering valuable frequent patterns based on RFM analysis without customer identification information. *Knowledge-Based Systems*, (61), 76-88.
- Jianying, M., Yongjian, F. & Yanguang, S. (2009). A neural networks-based clustering collaborative filtering algorithm in E-commerce recommendation system, Paper presented at the Web Information

Systems and Mining. *WISM 2009. International Conference*. IEEE, PP. 616-619.

Johnson, G. (2005). *Sales Forecasting for Pharmaceuticals: An Evidence Based Approach*. London. Scientific Publishing.

Khatami-Firuzabadi, M. & Sadaghiani, J. S. (2015). A developing model for clustering and ranking bank customers. *International Journal of Electronic Customer Relationship Management*, 9(1), 73-86.

Kolarovszki, P. Tengler, J. Majerčáková, M. (2016). The New Model of Customer Segmentation in Postal Enterprise. *Procedia - Social and Behavioral Sciences*, (230), 121-127.

Kotler, F. (2006). *Principles of Marketing*. (5, Ed., & Parsian, Trans.) Tehran, Tehran: Dabestan. (in Persian)

Li, D. Dai, W. & Tseng, W. (2011). A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. *Expert Systems with Applications*, 38(6), 7186-7191.

Maleki, M. Hajilu, Z. (2016). Identifying & Segmenting Key Customers for Prioritizing them Based on Lifetime Value using RFM Model (Case study: Internet customer of Qom Telecommunications Company). *Quarterly Journal of Business Management*, 2(6), 461-478. (in Persian)

Mehta, J. S. Gawande, A. (2015). A purpose of data mining in banking sector, International. *Journal of Advance Research in Computer Science and Management Studies*, 3(3), 407-411.

Ngai, E. Hu, Y. Wong, Y. Chen, Y. Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559-569.

Padashi, A. (2013). *Customer Segmentation and definition of the strategy in each sector Based on RFML Model*. Master of MBA Thesis, University of Guilan. (in Persian)

Patak, M., Lestacowa, H., Curdova, C. & Vlckova, V. (2014). The E-Pharmacy Customer Segmentation Based on the Perceived Importance of the Retention Support Tools. *Procedia - Social and Behavioral Sciences*, (150), 552-262.

- Pradeep, R. & Singh, S. (2010). A survey of clustering techniques. *International Journal of Computer Application*, 7(12).
- Prins, S. & Stegwee, R.A. (2000). *Zorgproducten en geïntegreerde infromatie systemen*. (in Dutch) Handboek sturen met zorgproducten, F3100-3, December.
- Rai, S. & Shubha, P. (2010). A Survey of Clustering Techniques. *International Journal of Computer Applications*, (1), 0915-8881.
- Ramos, S., Duarte, J., Duarte, F. & Vale, Z. (2015). A data-mining-based methodology to support MV electricity customers' characterization. *Energy and Buildings*, (91), 16-25.
- Ranjan, J. (2009). Data mining in pharma sector: benefits. *International journal of health care quality assurance*, 22(1), 92-82.
- Sadarina, P., Kothari, M. & Gondaliya, J. (2013). Implementing Data Mining Techniques for Marketing of Pharmaceutical Products. *International Journal of Computer Applications & Information Technology*, 2(1). (ISSN: 2278-7720).
- Saleabadi, S. (2011). Customers segmentation in the broadcast industry using data mining. *International Conference on Management And Economics Industri engineering*, (6), 1-20. (in Persian)
- Salehabadi, S. (2015). Classification of pharmaceutical distribution industry customers Based on data mining. *International conference on management economics and industrial engineering*. Tehran. [http://www.civilica.com/Paper-ICMEI01-ICMEI01\\_383.html](http://www.civilica.com/Paper-ICMEI01-ICMEI01_383.html).
- Senthilkumaran, U., Manikdan, N. & Senthilkumar, N. (2016). Role of data mining on pharmaceutical industry 883lanni. *International Journal of Pharmacy & Technology*, 8(3), 16100-16106.
- Seyed Hosseini, M., Maleki, A. & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to asses the customer loyalty. *Expert System with Applications*, (37), 5259-5264. (in Persian)
- Soeini, A. & Fathalizadeh, E. (2012). Customer segmentation based onmodified RFM model in insurance industry. *International Journal of Modeling and Optimization*, 25, 101-104.
- Tsiptsis, K. & Chorianopoulos, A. (2009). *Data mining techniques in CRM . Inside customer segmentation*. John Wiley & Sons, Chichester.

- Velmurugan, T. & Santhanam, T. (2011). A survey of partition based clustering algorithms in data mining: An experimental approach. *Information Technology Journal*, (3), 478-484.
- Wei, J., Lin, Sh. & Weng, Ch. (2012). A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Systems with Applications*, (39), 5529-5533.
- Zare Hosseini, Z. & Mohammadzadeh, M. (2016). Knowledge discovery from patients' behavior via clustering-classification algorithms based on weighted eRFM and CLV model: An empirical study in public health care services. *Iranian Journal of Pharmaceutical Research*, 15 (1), 355-367.

