

ارائه یک الگو ترکیبی داده کاوی با استفاده از قواعد انجمنی و خوشه بندی برای تعیین استراتژی تخفیف دهی، مطالعه موردی شرکت پخش پگاه

نویسندگان: آمنه خدیورا^۱ و پروشات حامدی^{۲*}

۱. استادیار گروه مدیریت بازرگانی دانشگاه الزهرا.

۲. کارشناس ارشد مدیریت فناوری اطلاعات دانشگاه الزهرا.

*Email:hamedi.p@gmail.com

چکیده

در اکثر شرکت های فروش و پخش، مسئله ترفیعات فروش از جمله مباحث مهم در بازاریابی و فروش است. در این میان مهم ترین دغدغه بازاریابان این است که مناسب ترین سیاست ها را برای پیشبرد فروش شرکت ها انتخاب کنند. از جمله استراتژی های پیشبرد فروش، سیاست تخفیف دهی به مشتریان است. استفاده از یک سیاست تخفیف دهی یکسان برای همه مشتریان و بر روی تمامی کالاها، میزان موفقیت تخفیف را کاهش می دهد، سیاست گذاری برای تخفیف دهی نیازمند استفاده از الگوی است که قادر باشد مناسب ترین نوع تخفیف را به انواع مشتریان در خریدهای متفاوت آنها پیشنهاد دهد. این تحقیق از جهت هدف، کاربردی و به لحاظ روش، توصیفی می باشد و در آن از تکنیک های داده کاوی، به عنوان ابزاری برای ایجاد یک الگو سیاست گذاری تخفیف، برای تدوین الگو سیاست های تخفیف دهی مناسب به مشتریان گروه های مختلف پرداخته شد. در این تحقیق از الگوریتم های خوشه بندی و قواعد انجمنی برای تدوین الگو سیاست گذاری تخفیف دهی مناسب به مشتریان در یک شرکت پخش مواد غذایی استفاده شده است.

به منظور خوشه بندی مشتریان و دستیابی به تعداد خوشه بهینه از معیار RFM و الگوریتم K-Means بهره گرفته شده است و سپس با کمک الگوریتم Apriori قواعد انجمنی، به استخراج قواعد موجود در هر خوشه برای کالاهای مختلف پرداخته شد. در نهایت با توجه به مشخصات مشتریان هر خوشه، استراتژی های مناسب تخفیف به هر گروه در نظر گرفته شد. نتایج این تحقیق نشان می دهد بهترین تعداد خوشه برای مشتریان این شرکت هشت عدد می باشد. مشتریان شرکت در دو گروه مشتریان برتر و نامطمئن قرار گرفتند. استراتژی های تخفیف دهی برای هر گروه بر اساس تجزیه و تحلیل نتایج به دست آمده از قواعد انجمنی و مصاحبه با خبرگان این شرکت استخراج گردید و در نهایت برای خریدهای مختلف هر گروه سیاست تخفیف دهی خاصی تدوین گردیده است.

کلید واژه ها: داده کاوی، قواعد انجمنی، الگوریتم کای میانگین، مشتری، تخفیف دهی، RFM، الگوریتم Apriori

راهبردهای
بازرگانی

(دانشور و نشر)

Business
Strategies

- دریافت مقاله: ۹۳/۲/۱۲
- پذیرش مقاله: ۹۴/۳/۹

*Journal of
Business Strategies
Shahed University
Twenty-Second Year
No.05
Spring & Summer
2015*

دو فصلنامه علمی - پژوهشی
دانشگاه شاهد
سال بیست و دوم - دوره جدید
شماره ۵
بهار و تابستان ۱۳۹۴

مقدمه

در طی چند سال اخیر تعامل شرکت ها با مشتریانشان به طور قابل توجهی تغییر کرده است، به طوری که تداوم کسب و کار با مشتری تضمین بلند مدت ندارد. به همین دلیل برای موفقیت یک سازمان لازم است، مشتریان خود را به درستی درک کرده، نیازها و خواسته های آنها را پیش بینی کنند و با مجهز شدن به این اطلاعات، سلامت کاری خود را بهبود بخشند. رضایت مشتری میزان مطلوبیتی است که وی به خاطر ویژگی های مختلف کالا کسب می کند و منبع سودآوری و دلیلی برای ادامه فعالیت سازمان است [۱].

مشتریان سرمایه اصلی شرکت ها محسوب می شوند [۲]؛ مدیران بازاریابی تنها وقتی می توانند به موفقیت امید داشته باشند که بتوانند رفتارهای آینده مشتریان را پیش بینی کنند. برای این منظور شرکت ها می بایست داده هایی کامل از فعالیت های گذشته مشتریان خود داشته باشند، هدف اصلی از ایجاد پروفایل برای هر مشتری، آگاهی از فعالیت های گذشته مشتریان می باشد تا بتوان از آنها در داده کاوی استفاده کرد.

از جمله موارد استفاده داده کاوی در زمینه مدیریت ارتباط با مشتری CRM، می توان به تدوین راهبرد در مورد روش های ارتباط با مشتری، تشویق مشتریان به خرید بیشتر و وفادار کردن مشتری به خدمات و محصولات خود اشاره کرد. در بیشتر شرکت های تولیدی و یا شرکت های پخش مسئله ترفیعات فروش از جمله مباحث مهم در بازاریابی و فروش است. همانگونه که بیان شد، مهمترین دغدغه بازاریابان شرکت ها این است که مناسب ترین سیاست ها، برای پیشبرد فروش را انتخاب کنند. از جمله راهبرد های پیشبرد، سیاست تخفیف دهی به مشتریان است. سیاست گذاری برای تخفیف دهی نیازمند استفاده از الگوی است که قادر باشد مناسب ترین نوع تخفیف را به انواع مشتریان در خریدهای متفاوت آنها پیشنهاد دهد. هدف اصلی این تحقیق ارائه یک الگو هوشمند به منظور سیاست گذاری تخفیف دهی به مشتریان، به نحوی که قادر باشد مناسب ترین سیاست تخفیف دهی را به مشتریان و کالاهای مختلف در خریدهای متفاوت در شرکت کدبانو پیشنهاد دهد. لذا رویکرد داده کاوی می تواند به عنوان ابزاری برای ایجاد یک الگو هوشمند سیاست گذاری تخفیف استفاده شود. به نحوی که با خوشه بندی مشتریان

بر اساس شاخص های مناسب و برجسته گذاری آنها، ابتدا مشتریان را دسته بندی نمود و سپس با استفاده از قواعد انجمنی به تجزیه و تحلیل رفتار خرید هر یک از گروه های مشتریان پرداخت.

مروری بر ادبیات تحقیق تعاریف داده کاوی

به طور کلی تحلیل داده های توصیفی کامپیوتری، در مجموعه های بزرگ و پیچیده داده ها را داده کاوی نامیده اند [۳]. داده کاوی یک حوزه میان رشته ای با رشد سریع است که حوزه های مختلفی همچون پایگاه داده، آمار، یادگیری ماشین و سایر زمینه های مرتبط را با هم تلفیق کرده است تا اطلاعات و دانش ارزشمند نهفته در حجم بزرگی از داده ها را استخراج نماید^۱. داده کاوی، اکتشاف و تحلیل حجم زیادی از داده ها برای کشف الگوها و قواعد معنادار است. فرآیند داده کاوی گاهی کشف دانش نیز نامیده می شود [۴].

فرآیند داده کاوی (CRISP)

موفق ترین پروژه های داده کاوی، در چارچوب فرآیند استاندارد اجرا می شود که توسط یک تیم کاری در شرکت SPSS در قالب پروژه ای به نام فرآیند اجرای داده کاوی استا ندارد صنعتی CRISP-DM Cross ارائه شده است [۵]. بر طبق CRISP-DM یک پروژه داده کاوی معین، شامل چرخه حیات شش مرحله ای است (شکل ۱). در مرحله درک تجارت، اهداف تجاری، وضعیت، اهداف داده کاوی تعیین می شود و طرح پروژه تولید می گردد. در مرحله فهم داده، به جمع آوری و توصیف داده های اولیه، کشف داده، بررسی کیفیت داده پرداخته می شود. در مرحله آماده سازی داده، داده ها انتخاب و پاک سازی می شود و سپس به ساخت داده، جمع آوری داده و در نهایت طراحی ظاهر داده می پردازیم. در الگو سازی، تکنیک الگو سازی انتخاب می شود و طرح آزمون تولید می گردد و در نهایت الگو ساخته و مورد ارزیابی قرار می گیرد. مراحل ارزیابی بدین شرح می باشد: ارزیابی نتیجه ها، بازبینی فرآیند، تعیین گام بعدی، توسعه، طرح ریزی برای توسعه، طرح ریزی برای نگهداری و نظارت، تهیه گزارش پایانی، بازبینی پروژه.

^۱ ACM SIGKDD, 2006

اساس معیاری که قبلاً تعیین شده انجام می شود. لیو در سال ۱۹۶۸ نشان داد که تعداد حالت های مختلف برای خوشه بندی n شیء به k خوشه برابر است با:

$$NW(n, k) = \frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n$$

از عبارت فوق به راحتی می توان فهمید که حتی اگر k از قبل تعیین شده باشد، پیدا کردن بهترین خوشه بندی کار ساده ای نیست. اما کاربردهای زیادی وجود دارد که در آنها مقدار k از قبل تعیین نشده است و کاربر نیز دید روشنی از تعداد خوشه هایی که در داده ها می توان مشخص کرد، ندارد. بنابراین باید از معیاری برای تعیین تعداد خوشه بهینه استفاده شود.

متداول ترین شاخص های مورد استفاده برای ارزیابی خوشه ها، شاخص دان^۱ و شاخص مجموع خطای مربعی SSE می باشند که در این تحقیق از این معیارها برای خوشه بندی استفاده شده است.

الگوهای RFM و LRFM برای تعیین ابعاد خوشه بندی

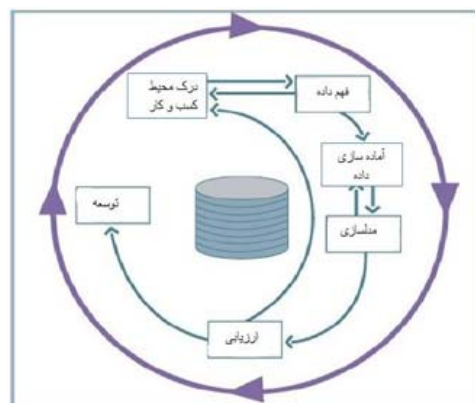
الگو RFM اولین بار توسط هوگس در سال ۱۹۹۴ معرفی گردید. وی برای تحلیل RFM از رفتار گذشته مشتری که به آسانی قابل پیگیری و دسترسی است، استفاده نمود. این الگو از سه بعد مربوط به داده های مبادلاتی مشتریان، برای تحلیل رفتار آنها استفاده می نماید. شاخص های این الگو به صورت زیر تعریف می شوند [۱۲].

شاخص تازگی مبادله (R)^۲ بر فاصله زمانی بین آخرین خرید صورت گرفته توسط مشتری تا پایان دوره خاص (پایان محدوده زمانی مورد بررسی) اشاره دارد. کم تر بودن این فاصله نشانگر بالا بودن ارزش این شاخص در الگو می باشد.

شاخص تعداد تکرار مبادله (F)^۳ بیانگر تعداد مبادلاتی

است که یک مشتری در یک دوره زمانی خاص انجام داده است. بیش تر بودن تعداد مبادلات، نشانگر بالا بودن ارزش این شاخص در الگو می باشد.

شاخص ارزش پولی مبادله (M)^۴ نشان دهنده مقدار پولی است که یک مشتری در یک دوره زمانی خاص جهت مبادلات، صرف نموده است. بیشتر بودن مقدار پول صرف



شکل ۱: فرآیند کریسپ

خوشه بندی

تحلیل خوشه ای که اصطلاح آن نخستین بار توسط تریان در سال ۱۹۳۹ به کار برده شد، یک تکنیک طبقه بندی برای تشکیل گروه های همگون در مجموعاً پیچیده ای از داده ها است [۶]. داده ها بر اساس اصل حداکثر کردن شباهت داخل گروه ها و حداقل کردن شباهت بین گروه ها، خوشه بندی می شوند. خوشه بندی یک روش متداول توصیفی است که در جستجوی تشخیص تعداد محدودی خوشه، برای توصیف داده ها است [۷]. بوسیله خوشه بندی می توانیم نحوه توزیع الگوها را مشخص نماییم. این تکنیک با روش های طبقه بندی Classification متفاوت است. طبقه بندی با تعداد معلوم و مشخصی از گروه ها ارتباط دارد، و هدف آن اختصاص مشاهده های جدید به یکی از این گروه ها است. اما تحلیل خوشه ای یک تکنیک ابتدائی است که به هیچ مفروضه ای درباره تعداد یا ساختار گروه ها متکی نیست. در تحلیل خوشه ای، عضویت گروهی برای همه مشاهده ها نامعلوم، و حتی تعداد گروه ها نیز نامشخص است [۸].

تقسیم مشتریان با ویژگی های معمول به گروه های کوچکتر می تواند مرجع مفیدی برای تصمیم گیری باشد. تاکنون الگوریتم K-means پر استفاده ترین تکنیک خوشه بندی برای گروه بندی مشتریان بوده است [۹]. در تحقیق های مشابه خوشه بندی، از شاخص RFM استفاده شده است [۱۰، ۱۱].

در بسیاری از روش های ارائه شده برای خوشه بندی، مقدار k (تعداد خوشه ها) باید توسط کاربر تعیین شود. در این روش ها جستجو برای یافتن بهترین تعداد خوشه بر

¹ Dunn Index

² Regency

³ Frequency

⁴ Monetary

همکاران مطرح شد [۱۵]. در واقع با داشتن مجموعه ای از تراکنش ها، یک قانون وابستگی به شکل $X \rightarrow Y$ بیان می شود (X و Y مجموعه ای از اقلام هستند). برای قواعد انجمنی دو پارامتر با عنوان های پشتیبانی^۲ و اطمینان^۳ معرفی می گردند. قوانینی انتخاب خواهند شد که مقادیر پشتیبانی و اطمینان آنها، از مقادیر حداقلی که توسط کاربر معین می گردد بالاتر باشند. لذا قوانینی که دارای مقادیر بالایی در پشتیبانی و اطمینان باشند، از اعتبار بیشتری برخوردار خواهند بود. پشتیبانی درصدی از جمعیت است که در یک قاعده، هم مقدم و هم تالی را دارند. به عبارتی درصدی از تراکنش ها که شامل همه موارد ظاهر شده در مقدم و تالی باشند.

$$\text{supp}(X \Rightarrow Y) = \text{supp}(X \cup Y) = \frac{|T(X \cup Y)|}{|T|}$$

اطمینان: در یک جمعیت مورد بررسی، درصدی از موارد است که وقتی مقدم قاعده در آنها ظاهر شده است، تالی نیز در آنها وجود دارد. مقدار اطمینان میزان وابستگی یک کالای خاص را به دیگری بیان می کند (بیانگر آن است که در صورت رخ دادن طرف اول قانون، طرف دوم قانون را در پی خواهد داشت).

$$\text{conf}(X \Rightarrow Y) = \frac{|T(X \cup Y)|}{|T(X)|}$$

به طوریکه $I = \{I_1, I_2, I_3, \dots, I_m\}$ مجموعه اقلام خریداری شده می باشد و T هر زیرمجموعه ای از I می باشد که از آن به عنوان تراکنش یاد می شود.

$X \quad I, Y \quad I$
 اشتراک X و Y تهی می باشد.

یکی از پایه ای ترین الگوریتم های کشف قوانین وابستگی الگوریتم Apriori می باشد؛ این الگوریتم که یکی از شناخته شده ترین الگوریتم های کشف قوانین وابستگی می باشد، توسط اگراوال و همکاران معرفی شد [۱۵]. روال کلی الگوریتم Apriori به شرح زیر می باشد:

دو مرحله اصلی تشکیل دهنده روال کلی این الگوریتم هستند. در یک مرحله، مجموعه کاندیداهای k عضوی C_k از پیوند L_{k-1} با خودش به وجود می آید و سپس در مرحله بعد، تعداد تکرار هر کدام از مجموعه های C_k محاسبه شده و بعد از آن براساس معیار حداقل تکرار قابل قبول، مجموعه هایی که تعداد تکرار کمی دارند حذف می شوند

شده، بیانگر بالا بودن ارزش این شاخص در الگو می باشد. تلاش های متعددی در زمینه بهبود متغیرهای بخش بندی صورت گرفته است و یکی از جدیدترین آن ها در واقع توسعه ای از الگو رایج RFM است، معیار بخش بندی LRFM، با یک رویکرد رفتار محور سعی بر تقسیم بندی مشتریان دارد که در آن علاوه بر شاخص های تأخر، تکرار و هزینه مالی، طول مدت همکاری مشتریان با سازمان نیز در قالب شاخص L مورد بررسی قرار می گیرد. مزیت این روش نسبت به RFM از آن جهت است که با افزودن شاخص طول مدت همکاری، بر رابطه بلند مدت با مشتریان تمرکز دارد و لذا ضمن بخش بندی مشتریان قادر به شناسایی و معرفی مشتریان وفادار نیز می باشد. معیار طول مدت همکاری Length. معرف تعداد روزهایی است که از اولین خرید مشتری و آغاز همکاری او با سازمان می گذرد.

قوانین انجمنی^۱

فرآیند کشف قوانین وابستگی، یکی از رویکردهای مهم در علم نوین داده کاوی برای یافتن قواعد و الگوها در پایگاه داده [۱۳] است که بسیار مورد توجه تحقیقگران قرار گرفته است. یکی از کاربردی ترین حالات تحلیل وابستگی ها "تجزیه و تحلیل سبد بازار" می باشد که در آن هدف یافتن کالاهایی است که معمولاً به طور همزمان خریداری می شوند. پیشرفت فناوری، فروشگاه های خرده فروش را قادر ساخته است تا حجم زیادی از داده های خرید مشتریان (که از آن به عنوان سبد خرید یاد می شود) را جمع آوری و ذخیره نماید. هر مشتری خرید مجزایی را در مقادیر مختلف و زمان های متفاوت انجام می دهد و داده های موجود در سبد بازار، نشان دهنده خرید مشتری در یک زمان خاص است. با تجزیه و تحلیل سبد بازار خرده فروشان، رفتار خرید مشتریان را پیش بینی کنند. این کار به آنها کمک می کند تا بتوانند کالاهای خود را بهتر ساماندهی کرده و چیدمان بهتری از محصولات خود داشته باشند و از این طریق سودآوری خود را افزایش دهند [۱۴]. هدف اصلی داده کاوی در پیدا کردن وابستگی، یافتن قانون های محکم و قابل توجه است. قوانین وابستگی اولین بار در سال ۱۹۹۳ برای مقادیر صفر و یک توسط اگراوال و

² Support

³ Confidence

¹ Association Rules

دوم روش تحلیل RFM توسعه داده شده با یک پارامتر اضافی - تعداد اقلام^۴ - استفاده شد که نتایج رویکرد دوم نشان دهنده آن بود که اضافه کردن پارامتر تعداد اقلام در نتایج بخش بندی مشتریان بی تأثیر است بنابراین CLV^۵ بر پایه روش RFM وزن دهی شده برای هر بخش محاسبه شد [۱۹].

سال ۲۰۱۱ مقاله ای با عنوان "یک روش دو مرحله ای خوشه بندی برای تحلیل مشخصات مشتریان به منظور مدیریت متمایز مشتریان"، توسط چیانگ لی، لی دای و تینگ تسنگ ارائه شد [۲۰]. این مطالعه از الگو LRFM که شامل چهار بعد - طول مدت همکاری، تأخر، تکرار خرید و مبلغ خرید - می باشد، استفاده کرد. از این تحلیل خوشه بندی برای طبقه بندی مشتریان جهت تنظیم راهبرد های بازایابی متفاوت استفاده شد. نتایج بدست آمده از کسب و کار نساجی نشان می دهد که گروه های مشتریان تفاوت های آماری معنی داری دارند و این مطالعه سازمان را در شناخت بهتر مشتریان، برای تعیین راهبرد های بازاریابی بخشی، یاری می کند.

در تحقیق "شناسایی و کشف الگوهای خرید مشتریان لوازم خانگی از طریق داده کاوی" سعی شده است تا روابط پنهان میان داده های حاصل از فروش محصولات خانگی با استفاده از تکنیک قوانین انجمنی کشف شود. نتایج حاصل از این روش کمک شایان توجهی در شناخت رفتار خرید مشتریان لوازم خانگی، طراحی و تدوین برنامه های بازاریابی جدید در حوزه لوازم خانگی، هدفمند ساختن بودجه های تبلیغاتی، امکان سنجی بازار برای تولید و عرضه محصول جدید و بررسی و شناخت قوت و ضعف محصولات و خدمات خواهد نمود [۲۱].

پژش "آنالیز خوشه بندی با استفاده از رویکرد داده کاوی برای توسعه متدولوژی CRM به منظور ارزیابی وفاداری مشتری" با کمک الگو توسعه یافته RFM و تعیین مقدار بهینه K با کمک الگوریتم K-means، به بخش بندی وفاداری محصول مشتری پرداختند؛ نتایج این تحقیق قابلیت شگرف سازمان در ارزیابی وفاداری مشتری با استفاده از این یافته ها را در استراتژی های بازاریابی نشان می دهد [۲۲].

و مجموعه های باقیمانده در این مرحله L_k را تشکیل خواهند داد. دو مرحله بیان شده، آنقدر تکرار می شوند، تا L_k جدیدی به دست نیاید. بنابراین شرط الگوریتم تهی بودن L_k خواهد بود. این الگوریتم، برای محاسبه اقلام کاندیدا، همه تراکنش ها را بررسی می کند و بنابراین نیازمند زمان زیادی است [۱۴].

پیشینه تحقیق

در تحقیقی مشابه که توسط غضنفری و همکاران در ۱۳۸۸ ارائه شد، هدف تحقیق ارائه الگوی برای بخش بندی کشورها براساس تازگی^۱، ارزش پولی^۲ و تکرار^۳ صادرات میوه های خوراکی از کشور جمهوری اسلامی ایران به سایر کشورها بود، پس از محاسبه متغیرها، با استفاده از الگوریتم های خوشه بندی (Kmeans و Fuzzy kmeans) کشورها خوشه بندی شده اند و نتایج آن از نظر کیفیت براساس سه معیار مختلف سنجیده شده است [۱۶].

در تحقیق بابایی و همکاران ۲۰۱۱، با تجزیه و تحلیل رفتار مشتریان، تعیین وزن و اهمیت هر یک از شاخه های به کار رفته در تحلیل (RFM) - که با استفاده از فرآیند تجزیه و تحلیل سلسله مراتبی به دست آمده اند - و نیز رتبه بندی و مقایسه مشتریان، به مقایسه مشتریان مهمان بانک ملت و همچنین تسهیل در تدوین استراتژی های بازاریابی در جهت برخورد با هر یک از این دسته ها پرداخته شده است [۱۷].

هنری چان و همکاران [۱۸]، در سال ۲۰۱۱، در تایوان در تحقیق خود به بخش بندی مشتریان به کلاس های مختلف بر پایه داده های RFM آنها و ارائه لیست قیمت محصولات مطابق با کلاسی که هر مشتری در آن قرار داده شده است و در نهایت در نظر گرفتن ترفیعاتی برای مشتریان با ارزش پرداختند.

در تحقیقی با عنوان "ارزیابی ارزش طول عمر مشتری بر پایه تحلیل RFM رفتار خرید مشتریان با استفاده از ارزش طول عمر مشتری به بخش بندی مشتریان یک انستیتو زیبایی و سلامتی پرداخته شد. دو رویکرد مورد بررسی قرار گرفت: در رویکرد اول از روش تحلیل بازاریابی RFM برای بخش بندی مشتریان و در رویکرد

⁴ Count Item

⁵ Customer Lifetime Value

¹ regency

² Mote nary

³ Frequency

روش شناسی تحقیق

نوع تحقیق کاربردی می باشد.

طبقه بندی تحقیق بر مبنای روش

غالب مطالعات تحقیقی یک روش یا راهبرد که شامل رویه های مشترک خاصی است را نشان می دهد. با توجه به این امر که این تحقیق با توصیف و تفسیر شرایط و روابط موجود بین کالاها و مشتریان، به دسته بندی مشتریان صنعت تولیدی می پردازد و مجموعه ای ساختمان و منظمی از داده ها را جمع آوری می کند، از نوع تحقیق توصیفی می باشد.

طبقه بندی تحقیق بر مبنای نوع داده ها

براساس این طبقه بندی به دو دسته تحقیق می توان اشاره کرد:

تحقیق های کمی

تحقیق های کیفی

با توجه به اینکه تمام داده های این تحقیق در قالب اعداد و ارقام هستند، تحقیق حاضر از نوع کمی است.

اجرای داده کاوی

پیشن پردازش و پاکسازی داده ها

داده های مربوط از پایگاه داده شرکت پخش مورد مطالعه تهیه شده است، که شامل ۵۴۴۹۹۸ رکورد می باشد و اطلاعات مربوط، ۱۳۲۱۴ مشتری را شامل می شود؛ نمونه آماری، متشکل از داده های مربوط بخش فروش شرکت کدبانو از تاریخ ۱۳۹۱/۰۲/۰۱ تا ۱۳۹۱/۰۸/۳۰ می باشد.

در این تحقیق به منظور بخش بندی مشتریان رویکردهای LRFM و RFM جداگانه با کای میانگین تلفیق شدند تا نتایج آنها با هم مقایسه شود. شاخص های مورد استفاده در این دو الگو درباره داده های در دسترس به شرح جدول ۱ تعریف شدند. با توجه به اینکه محدوده مقادیر شاخص ها متفاوت است، برای جلوگیری از تأثیر آن بر نتایج خوشه بندی لازم بود مقادیر نرمال سازی شوند.

جدول ۱: شاخص های RFM و LRFM در این تحقیق

طول مدت همکاری (L)	تعداد روزهایی که از اولین خرید مشتری در طول تاریخ همکاری تا تاریخ ۱۳۹۱/۰۸/۳۰ می گذرد.
تأخر خرید (R)	تعداد روزهایی که از آخرین خرید تا تاریخ ۱۳۹۱/۰۸/۳۰ می گذرد.
تکرار خرید (F)	تعداد فاکتورهای خرید مشتری در بازه ۱۳۹۱/۰۲/۰۱ تا ۱۳۹۱/۰۸/۳۰.
مبلغ خرید (M)	جمع کل مبالغ فاکتورهای مشتری در بازه ۱۳۹۱/۰۲/۰۱ تا ۱۳۹۱/۰۸/۳۰.

هدف از این تحقیق، کاوش در پایگاه داده محصولات یک شرکت پخش مواد غذایی، با استفاده از روش های داده کاوی، به منظور کشف الگوهای حاکم بر خرید و فروش محصولات و نهایتاً کشف دانش از این الگوها و ارائه قواعد استنتاج شده از این الگوها، به منظور کمک به مدیران در تصمیم گیری در مورد نحوه و مقدار بهینه اختصاص تخفیف به مشتریان می باشد. نوع تحقیق از نقطه نظر هدف کاربردی و از نقطه نظر داده، توصیفی می باشد. پرسش های اصلی این تحقیق به شرح زیر می باشد:

- الگو هوشمند برای تبیین سیاست های تخفیف دهی به مشتریان دارای چه ویژگی هایی است؟
 - مناسب ترین الگوریتم برای خوشه بندی مشتریان شرکت مورد مطالعه کدام است؟
 - بهترین تعداد خوشه جهت بخش بندی مشتریان با استفاده از شاخص های بدست آمده چه تعداد می باشد؟
 - مشتریان هر خوشه چه ویژگی هایی دارند؟
 - چه محصولاتی بیشتر با هم به فروش می روند؟
 - مشتریان هر خوشه به خرید چه محصولاتی تمایل دارند؟ و چه سیاست تخفیفی برای هر خوشه می توان در نظر گرفت؟
 - چگونه می توان از خوشه بندی و تحلیل سبد خرید مشتری، برای اتخاذ راهبرد مناسب تخفیف دهی به مشتریان استفاده کرد؟
- پایه هر علمی روش شناخت آن است و اعتبار و ارزش قوانین هر علمی به روش شناختی مبتنی است که در آن علم به کار می رود. روش های تحقیق را با معیارهای مختلف دسته بندی می کنند، در ادامه به موارد به کار گرفته شده در این تحقیق می پردازیم:

طبقه بندی تحقیق بر مبنای هدف

از آنجایی که این تحقیق قصد دارد با ارتقاء سطح دانش دسته بندی مشتریان بر مبنای عامل های تأثیرگذار، به تعریف سیاست های مناسب در تخفیف دهی پردازد، از

۲، شاخص ارزیابی دان $K=2$ با مقدار $1/39$ را بهینه ترین خوشه بندی تعیین نمود.

برای اینکه از تعداد خوشه بهینه اطمینان حاصل کنیم و بتوانیم خوشه بندی های خوبی داشته باشیم از شاخص SSE برای ارزیابی تعداد خوشه ها استفاده نمودیم (جدول ۳).

از بین تمام خطاها، مقدار مینیمم را به عنوان تعداد خوشه بهینه انتخاب می کنیم. با توجه به نتایج مندرج در جدول فوق معیار مجموع خطای مربعی، ۱۰ را تعداد خوشه بهینه تعیین کرده است. همان طور که مشاهده شد تفاوت تعداد خوشه های بهینه بدست آمده از شاخص ها زیاد است؛ همانطور که قبلاً اشاره شد، متغیر L نشان دهنده طول مدت همکاری مشتریان با شرکت است، از آنجا که داده های در اختیار ما از تاریخ $1391/02/01$ تا $1391/08/30$ می باشد، و مدت زمان همکاری افراد مختلف به هم نزدیک می باشد در نتیجه L حضور مؤثری در خوشه بندی ما نداشته است. در نتیجه تمام مراحل فوق را برای خوشه بندی بر مبنای RFM نیز انجام دادیم که به جهت اختصار از ارائه آنها خودداری شده است.

برای نرمال سازی روش های مختلفی وجود دارد. در این تحقیق به منظور نرمال سازی از امتیازدهی با استفاده از نرم افزار کلمنتاین استفاده شده است که به کلیه چهار شاخص، عددی بین یک تا پنج تخصیص می دهد، بنابراین به جای مقادیر معیارهای RFM و LRFM، امتیازات LRFM. Score آن ها به عنوان ورودی الگوریتم خوشه بندی در نظر گرفته شد.

خوشه بندی با معیار RFM و LRFM با استفاده از الگوریتم کای میانگین

ابتدا با الگوریتم K-Means بر اساس معیار LRFM، خوشه بندی برای ۲ تا ۱۰ خوشه انجام گرفت و سپس برای تعیین تعداد خوشه بهینه از دو شاخص ارزیابی دان و SSE استفاده شد. معیار دان را برای تمامی خوشه ها در هر خوشه بندی محاسبه کرده و برای مقایسه خوشه بندی های مختلف با تعداد خوشه های متفاوت، برای هر تعداد خوشه، ماکزیمم مقدار این معیار را برای هر خوشه بندی به دست آوردیم. برای شناخت بهترین خوشه بندی از بین این مقادیر بهترین حالت را انتخاب می کنیم یعنی مینیمم می گیریم. خوشه بندی که مقدار مینیمم به آن تعلق دارد، خوشه بندی بهینه خواهد بود؛ با توجه به اعداد بدست آمده در جدول

جدول ۲: محاسبه شاخص تعریف شده به ازای تعداد خوشه های مختلف

تعداد خوشه ها	محاسبه میزان d_{max} در هر خوشه بندی
$2K=$	$1/39d_{max}=$
$3K=$	$2/02d_{max}=$
$4K=$	$1/67d_{max}=$
$5K=$	$2/18d_{max}=$
$6K=$	$2/08d_{max}=$
$7K=$	$1/62d_{max}=$
$8K=$	$1/73d_{max}=$
$9K=$	$1/62d_{max}=$
$10K=$	$1/53d_{max}=$

جدول ۳: مقادیر شاخص ارزیابی SSE همه خوشه ها

	K=2	K=3	K=4	K=5	K=6
SSE	59553/24	37454/57	41317/03	45924/88	50709/01
	K=7	K=8	K=9	K=10	
SSE	40212/23	32806/79	95077/68	29015/15	

جدول ۴: مشخصات هر خوشه

تعداد رکوردها	میانگین M	میانگین F	میانگین R
۲۵۷۴	۵(٪۶۴/۵۷)	۵(٪۹۹/۱۵)	۴(٪۲۹/۲۲)
۲۰۴۷	۱(٪۹۴/۴۸)	۱(٪۹۸/۷۳)	۱(٪۵۷/۱۶)
۱۹۲۱	۳(٪۷۰/۰۷)	۳(٪۹۸/۹۱)	۲(٪۲۶/۸۱)
۲۶۳۹	۴(٪۵۱/۴۶)	۴(٪۹۹/۲۴)	۴(٪۲۷/۲۵)
۱۰۶۵	۲(٪۷۲/۳۹)	۲(٪۶۸/۵۴)	۴(٪۵۸/۶۹)
۹۳۳	۲(٪۴۸/۷۷)	۲(٪۴۴/۹۱)	۳(٪۱۰۰)
۹۹۷	۲(٪۶۸/۶۱)	۲(٪۶۸)	۲(٪۱۰۰)
۱۰۳۸	۲(٪۶۹/۱۷)	۲(٪۷۲/۷۴)	۱(٪۱۰۰)

تجزیه و تحلیل نتایج حاصل از داده کاوی به منظور سیاست‌گذاری تخفیف‌دهی

به منظور تجزیه و تحلیل نتایج، خروجی های به دست آمده از داده کاوی با خبرگان حوزه بازاریابی و فروش شرکت به بحث و بررسی گذاشته شده است [۲۴]. با توجه به امتیازات متغیرهای تأخر خرید، تکرار خرید و ارزش پولی هشت حالت برای مشتریان به وجود می آید که در جدول ۶ تشریح شده است.

با توجه به جدول فوق مشتریان گروه اول، مشتریان خوب و وفادار شرکت می باشند که دارای تکرار خرید بالایی هستند، در واقع رفتار این گروه از مشتریان برای شرکت معنا دار و قابل اعتنا است چرا که مشتریان اصلی شرکت به حساب می آیند؛ هدف از هرگونه تخفیف دهی به این گروه از مشتریان افزایش ارزش مالی خرید (M) آنها از طریق استراتژی های تخفیف‌دهی خواهد بود. این هدف با استراتژی کلی توسعه فروش سازگار است. برای تدوین استراتژی های دقیق مربوط به این گروه به قواعد انجمنی و الگوهای خرید آنها که برای ما معنی دار است مراجعه می کنیم [۲۵]. همانطور که در جدول ۶ مشاهده می شود، گروه اول خود به دو دسته تقسیم می شوند.

دسته اول، مشتریانی با تکرار خرید و ارزش پولی بالایی هستند، که مشتریان برتر محسوب می شوند، سیاست‌های مناسب برای مشتریان برتر به شرح زیر می‌باشد:

سیاست یک: با مشاهده قواعد استخراج شده مربوط به این دسته، این نتیجه حاصل می شود که نباید به کالاهای موجود در قواعد با درجه پشتیبان بالای ۳۰ درصد، تخفیفی تعلق گیرد، چون در بیشتر از ۳۰ درصد مواقع این کالاها

با توجه به نتایج حاصل از شاخص دان، تعداد خوشه بهینه ۸ می باشد. نتایج حاصل از شاخص SSE، تعداد خوشه بهینه را ۹ نشان می دهد. برای تعیین تعداد خوشه بهینه، مقدار بدست آمده از هر شاخص را برای ۹ خوشه و ۸ خوشه مقایسه کرده تا مشخص شود پاسخ هر شاخص ارزیابی چقدر مورد تأیید شاخص دیگر است و در نهایت ۸ خوشه به عنوان تعداد بهینه انتخاب شد. مشخصات هر خوشه در جدول زیر نشان داده شده است (جدول ۴).

پیاده‌سازی قوانین انجمنی

در این مرحله به تجزیه و تحلیل رفتار مشتریان هر خوشه پرداخته می شود. پس از ایجاد تغییرات لازم در داده ها به منظور تحلیل سبد خرید مشتری، برای اینکه رابطه بین کالاهای خریداری شده را مشخص کنیم، از نرم افزار Clementine و الگوریتم Apriori استفاده کردیم [۲۳]. استفاده از این الگوریتم به جهت سرعت بالای پاسخ دهی مورد توصیه تحقیقگران بوده است. در آغاز مقدار حداقل پشتیبان مقدم و حداقل اطمینان را مشخص می کنیم، افزایش این مقادیر منجر به ایجاد قوانین کمتر ولی معتبرتر می شود. از خروجی های خوشه بندی الگو کای میانگین - که در بخش قبل هشت خوشه تعیین شد - به عنوان ورودی الگوریتم Apriori استفاده شده است؛ برای این کار برای هر خوشه یک جدول شامل تمام فاکتورهای مربوط به مشتریان همان خوشه تشکیل شده و اطلاعات هر خوشه را با الگوریتم Apriori الگو شده است.

نتایج حاصل از تحلیل سبد خرید مشتریان هر خوشه و قواعد استخراج شده در هر خوشه به شرح زیر است (جدول ۵):

جدول ۵: قواعد حاصل برای هر خوشه

خوشه	قاعده	تالی	مقدم	پشتیبان	اطمینان
۱	۱	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس مایونز شیشه خانواده ○ سس سالاد فرانسوی ○ سس کچاپ موشکی کوچک 	۲۲/۳۲۵	۹۰/۷۷۴
۱	۲	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم ○ سس کچاپ موشکی کوچک 	۲۲/۸۷۲	۹۰/۷۵۴
۱	۳	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس کچاپ موشکی متوسط ○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم 	۲۰/۲۹۵	۹۰/۲۲۸
۱	۴	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس مایونز شیشه خانواده ○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم 	۲۵/۱۳۵	۸۹/۶۶۴
۱	۵	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس کچاپ موشکی کوچک 	۳۳/۹۶۶	۸۹/۱۳۳
۱	۶	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس مایونز فشاری کوچک ○ سس سالاد فرانسوی 	۲۴/۶۹	۸۸/۹۴۶
۱	۷	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس کچاپ موشکی متوسط ○ سس سالاد فرانسوی 	۲۸/۴۳۲	۸۸/۷۶۱
۱	۸	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس مایونز شیشه کوچک ○ سس سالاد فرانسوی 	۲۲/۲۸۲	۸۸/۲۴۶
۱	۹	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس مایونز شیشه خانواده ○ سس سالاد فرانسوی 	۳۲/۲۲۴	۸۸/۰۵۹
۱	۱۰	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم 	۳۳/۰۸۵	۸۷/۹۸۵
۱	۱۱	سس کچاپ موشکی کوچک	<ul style="list-style-type: none"> ○ سس کچاپ تند کوچک ○ سس سالاد هزار جزیره 	۲۲/۱۰۹	۸۶/۶۴
۱	۱۲	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی 	۵۲/۳۴۶	۸۵/۶۳۸
۲	۱	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی 	۲۳/۶۷	۶۳/۰۹۷
۲	۲	سس کچاپ موشکی کوچک	<ul style="list-style-type: none"> ○ سس کچاپ تند کوچک 	۱۰/۷۲۵	۶۰/۳۵۲
۳	۱	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس کچاپ موشکی کوچک 	۱۳/۱۳۵	۷۷/۲۸۹
۳	۲	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس مایونز شیشه خانواده 	۱۱/۳۶۳	۷۵/۰۱۸
۴	۱	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس کچاپ موشکی کوچک 	۱۶/۸۸	۷۹/۵۴۶
۴	۲	سس سالاد هزار جزیره	<ul style="list-style-type: none"> ○ سس سالاد فرانسوی ○ سس مایونز شیشه خانواده ○ سس مایونز شیشه نیم 	۱۰/۱۵۴	۷۹/۲۲۵

۷۸/۷۴۷	۱۵/۴۰۴	○ سس کچاپ موشکی متوسط ○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۳	۴
۷۸/۲۰۸	۱۲/۶۵۶	○ سس مایونز فشاری کوچک ○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۴	۴
۷۷/۰۹۸	۱۵/۹۴۶	○ سس سالاد فرانسوی ○ سس مایونز شیشه خانواده	سس سالاد هزار جزیره	۵	۴
۷۶/۰۰۴	۱۷/۱۷۷	○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم	سس سالاد هزار جزیره	۶	۴
۶۵/۴۸۱	۲۷/۴۰۶	○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۱	۴
۶۵/۴۸۱	۲۷/۴۰۶	○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۱	۵
۶۲/۸۱۲	۱۰/۲۹۲	○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم	سس سالاد هزار جزیره	۱	۶
۶۲/۳۷۵	۲۸/۰۹۸	○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۲	۶
۶۱/۱۶۳	۱۰/۰۳۵	○ سس کچاپ تند کوچک	سس کچاپ موشکی کوچک	۳	۶
۶۷/۷۰۵	۲۶/۳۶۷	○ سس سالاد فرانسوی	سس سالاد هزار جزیره	۱	۷
۷۴/۳۰۸	۱۰/۸۳۵	○ سس سالاد فرانسوی ○ سس کچاپ موشکی کوچک	سس سالاد هزار جزیره	۱	۸
۷۰/۷۱۴	۱۱/۹۹۱	○ سس سالاد فرانسوی ○ سس مایونز شیشه نیم	سس سالاد هزار جزیره	۲	۸
۷۰/۵۳۱	۱۳/۲۹۸	○ سس مایونز شیشه کوچک ○ سس مایونز شیشه خانواده	سس مایونز شیشه نیم	۳	۸

جدول ۶: برجسب گذاری خوشه ها بر اساس شاخص RFM

شماره گروه	برجسب هر گروه	ارزش پولی	تکرار خرید	تاخر خرید	تعداد حالات
گروه یک	مشتریان برتر (دسته یک)	M↑	F↑	R↑	۱
		M↑	F↑	R↓	۲
گروه دو	مشتریان تکرار شونده (دسته دو)	M↓	F↑	R↑	۳
		M↓	F↑	R↓	۴
گروه دو	مشتریان خرج کننده (دسته سه)	M↑	F↓	R↑	۵
		M↑	F↓	R↓	۶
گروه دو	مشتریان نا مطمئن (دسته چهار)	M↓	F↓	R↑	۷
		M↓	F↓	R↓	۸

هزار جزیره، سس مایونز شیشه ساین خانواده و سس مایونز شیشه نیم کیلویی هستند. از میان این محصولات سس سالاد فرانسوی و سس سالاد هزار جزیره درجه پشتیبان بالاتری دارند و از محصولاتی به شمار می روند که در اکثر مواقع به فروش می روند و شرکت در فروش آنها مشکلی ندارد؛ در نتیجه هیچ نیازی به در نظر گرفتن تخفیف روی

خریداری می شوند و نیازی به در نظر گرفتن ترفیع فروش برای این محصولات وجود ندارد. به عبارت دیگر نتایج نشان داده که در مورد این کالا سیاست عدم تخفیف دهی باید رعایت شود. در مورد این دسته از مشتریان شرکت، محصولاتی که نیاز به تخفیف ندارند پنج محصول سس سالاد فرانسوی، سس کچاپ موشکی کوچک، سس سالاد

این محصولات وجود ندارد.

سیاست دو: به منظور افزایش ارزش مالی در این خوشه ها به قواعدی که درجه اطمینانی بین ۸۵ تا ۹۵ درصد دارند و در آنها حداقل سه محصول از پنج محصول پر فروشی که در بالا نام برده شد، وجود دارد، تخفیف متوسطی برای بقیه کالاهای موجود در سبد خرید در نظر گرفته می شود. این بدان معنی است که در ۸۵ درصد مواقع اگر فرد سه کالا از پنج کالای مذکور را خریداری نماید، کالاهای دیگری نیز تهیه می کند. برای اینکه این درصد را افزایش دهیم و مشتری را حتماً به خرید کالاهایی که تمایل دارد ترغیب کنیم، با در نظر گرفتن تخفیفی متوسط برای سایر کالاها، با اعمال تشویقی نه چندان پر هزینه آنها را به خرید سایر کالاها (منظور کالاهایی که درصد اطمینان پایینی با این کالاها دارند) هدایت نموده تا درجه اطمینان را به صد درصد افزایش دهیم.

دسته دوم، مشتریان تکرار شونده می باشند که تکرار خرید بالائی دارند ولی ارزش مالی آنها پایین است. این بدان معنی است که این دسته از مشتریان مرتب در هر ویزیت، خریدی به اندازه نیازشان انجام می دهند. در این دسته استراتژی تخفیف نباید به گونه ای باشد که بیش از حجم مورد نیازشان خرید کنند چون این امر ممکن است منجر به کاهش تکرار خرید شود و در نهایت منجر به کاهش F خواهد شد و این مطلوب نیست. لذا در این دسته از طریق افزایش اقلام سبد خرید مشتری و با پیشنهاد کالاهای متنوع به آنها ارزش مالی خرید را افزایش می دهیم. با توجه به قواعد استخراج شده از خوشه های این دسته، کالاهایی که در قواعدی با درجه پشتیبان بیش از ۲۰ درصد قرار دارند را شناسائی کرده و اگر در دوره ای از سفارش مشتری، تعدادی از این اقلام را سفارش داد، بر روی سایر کالاها (بقیه کالاهای این گروه نه هرکالایی) به عبارت دیگر مشتری را با کالاهای محبوب این دسته به افزایش مبلغ خرید ترغیب می کنیم) تخفیفی در نظر گرفته شود تا مشتری به خرید آن کالاها ترغیب شود. بدین معنی که به هر کالایی با مشخصات فوق که سفارش داده نشده باشد، تخفیف تعلق می گیرد. مثلاً کالای X, Y, Z و W طبق قاعده ای در بیش از ۲۰ درصد موارد با هم خریداری می شوند؛ حال مشتری در این دوره فقط کالاهای X و Y را خریداری کرده است، باید برای کالاهای Z و W تخفیفی برای این مشتری در نظر بگیریم

تا احتمال فروش این کالاها را زیاد کنیم و بدین ترتیب ارزش مالی این دسته از مشتریان را بهبود ببخشیم. مشتریان گروه دوم: بعد از بررسی مشخصات افراد در این گروه و گرفتن نظر مدیران بخش فروش و بازاریابی شرکت، این نتیجه حاصل شد که با توجه به رفتار و الگوی خرید این مشتریان که به صورت پراکنده خرید می کنند، نمی توان به قواعد استخراج شده از این خوشه ها استناد کرد و این قواعد اعتبار بالائی برای استفاده در تعیین سیاست تخفیف دهی ندارند و بهتر است به نتایج خوشه بندی استناد کرد. بنا به نظر خبرگان شرکت، این افراد جزء مشتریان اصلی شرکت محسوب نمی شوند و از رقبای هم می کنند و هدف اصلی از سیاست گذاری تخفیف این مشتریان، تبدیل این مشتریان به مشتریان وفادار شرکت است، به نحوی که ارزش مالی خرید یا تعداد تکرار خرید آنها افزایش یابد و تبدیل به گروه اول شوند. به عبارت دیگر باید از استراتژی ایجاد جای پا استفاده شود.

از آنجا که مشتریان این گروه دارای تکرار خرید پایینی هستند، به عنوان یک سیاست کلی در مورد این مشتریان، باید تخفیفی برای مشتریانی که دوره ویزیت خود را در سفارش دهی رعایت می کنند قائل شد، به نحوی که متغیر F بهبود یابد. در این گروه نیز با دو دسته مختلف مشتریان مواجه هستیم.

دسته سوم مشتریانی هستند که ارزش و حجم خرید خوبی دارند ولی تعداد دفعات خرید آنها کم است، در نتیجه هدف افزایش تعداد تکرار خرید در این دسته می باشد. بنابراین ویزیت منظم برای آنها توصیه می شود و در صورت خرید منظم، تخفیف نقدی را به این افراد تخصیص می دهیم. در این دسته، از استراتژی تخفیف دهی به اقلام استفاده می کنیم تا از این طریق تنوع کالاهای خریداری شده افزایش پیدا کند. تخفیف اقلام بدان معناست که اگر مشتری از تعداد قلم کالای مدنظر بیشتر خریداری کرد، روی کل مبلغ خرید تخفیفی در نظر گرفته شود. که در این شرکت با نظر خبرگان و مدیران ارشد، به ازای هر قلم خرید اضافه بر ۷ قلم کالا، نیم درصد تخفیف روی کل فاکتور دریافت خواهند کرد (تا سقف ۱۲ قلم کالا)؛ یعنی اگر مشتریان این گروه در هر بار سفارش دهی از هفت قلم کالا بیشتر سفارش دهند، روی مبلغ کل فاکتور از نیم درصد تخفیف برخوردار خواهند شد. دسته چهارم، مشتریان خوبی نیستند و باید تلاش شود

با اختصاص دادن تخفیف ریالی روی کل خرید آنها، این مشتریان را از رقبا دور و به خود نزدیک کنیم تا در نهایت به مشتریان خوشه های دیگر تبدیل شوند، یعنی ارزش مالی آنها یا تکرار خرید افزایش یابد. بنابراین طبق نظر خبرگان روی مبلغ کل خرید آنها تخفیف ریالی پلکانی قائل می شویم.

پیشنهادهای کاربردی جامعه مورد مطالعه

داده کاوی یک فرآیند مستمر می باشد و در یک مرحله پایان نمی پذیرد، باید توجه داشت که در هر مرحله از داده کاوی، تحلیلگر به تنهایی کاوش را ادامه نمی دهد، بلکه باید در کلیه مراحل با خبرگان و کاربران در تقابل بوده و با دریافت نظرها و پیشنهادهای آنها، مراحل تحقیق را به سوی کشف دانش بیشتر هدایت نماید. که این موضوع در این تحقیق در نظر گرفته شده و در تمام مراحل با خبرگان تعامل خوبی صورت گرفت.

با توجه به تجزیه و تحلیل نتایج حاصل از خوشه بندی مشتریان شرکت پخش مواد غذایی مورد مطالعه، سیاست های تخفیف تدوین شده برای هر دسته که در بخش قبل توضیح داده شد را به شرح زیر می توان به این شرکت و محصولات آن تعمیم داد.

همانطور که گفته شد، مشتریان خوشه یک، سه و چهار مشتریان برتر ما هستند که در گروه اول و در دسته اول قرار می گیرند که مشتریانی با تکرار خرید و ارزش پولی بالایی هستند؛ با توجه به استراتژی اول به پنج محصول سس سالاد فرانسوی، سس کچاپ موشکی کوچک، سس سالاد هزار جزیره، سس مایونز شیشه سبک خانواده و سس مایونز شیشه نیم کیلویی که کالاهایی با درجه پشتیبان بالای ۳۰ درصد تخفیفی هستند، تخصیص داده نمی شود چون در بیشتر از ۳۰ درصد مواقع این کالاها خریداری می شوند و نیازی به در نظر گرفتن ترفیع فروش برای این محصولات وجود ندارد (به طور کلی کالاهای پرفروش محسوب می شوند)؛ با توجه به قاعده سه از خوشه اول که با درجه اطمینان ۹۹/۲۲ محصول سس سالاد فرانسوی، سس مایونز شیشه ای نیم کیلویی و سس سالاد هزار جزیره با هم خریداری می شوند برای سس کچاپ موشکی سبک متوسط تخفیف متوسطی در نظر گرفته خواهد شد. و همچنین از قواعد با درجه اطمینان بین ۵۰ تا ۸۰ درصد و درجه پشتیبان حداقل ۲۰ درصد، می توان استراتژی های

تخفیف زیر را برای شرکت در نظر گرفت.
اگر محصولات سس مایونز شیشه ای نیم کیلویی، سس مایونز شیشه ای خانواده و سس سالاد هزار جزیره با هم خریداری شوند، سس مایونز شیشه ای سبک کوچک شامل تخفیف این گروه می شود. در صورت وجود کالاهای سس کچاپ موشکی سبک کوچک، سس سالاد فرانسوی در سبد خرید مشتری، سس مایونز فشاری سبک کوچک، کالائی است که تخفیف به آن تعلق می گیرد. خرید سس مایونز شیشه ای نیم کیلویی، سس سالاد هزار جزیره و سس کچاپ موشکی سبک کوچک، سس مایونز شیشه ای سبک کوچک را شامل تخفیف می کند. سس مایونز شیشه ای نیم کیلویی، سس مایونز شیشه ای خانواده و سس سالاد فرانسوی در سبد خرید مشتری، کالای سس کچاپ موشکی سبک متوسط را در محصولات شامل تخفیف قرار می دهد. با توجه به دو قاعده حاصل، در صورت همراهی حداقل سه محصول از محصولات سس مایونز شیشه ای نیم کیلویی، سس مایونز شیشه ای خانواده، سس سالاد هزار جزیره، سس سالاد فرانسوی و سس کچاپ موشکی سبک کوچک، تخفیفی شامل سس کچاپ تند سبک کوچک می شود. خرید حداقل سه محصول از سس مایونز شیشه ای نیم کیلویی، سس مایونز شیشه ای سبک خانواده، سس سالاد هزار جزیره، سس سالاد فرانسوی، سس مایونز دبه ای را شامل تخفیف می کند. با توجه به استراتژی های فوق، محصولاتی که با توجه به قواعد انجمنی خوشه یک که در گروه اول قرار گرفته اند، شامل تخفیف شدند، سس مایونز شیشه ای سبک کوچک، سس مایونز فشاری سبک کوچک، سس کچاپ موشکی سبک متوسط، سس کچاپ تند سبک کوچک و سس مایونز دبه ای بودند.

نظر به اینکه مشتریان بقیه خوشه ها (خوشه دو، پنج، شش، هفت و هشت) در گروه دوم و دسته چهارم قرار می گیرند که از مشتریان بد و نامطمئن هستند، باید تلاش شود با اختصاص دادن تخفیف ریالی روی کل خرید آنها، این مشتریان را به خرید از شرکت تشویق و از رقبا دور نمانیم. همانطور که ذکر شد به دلیل پراکندگی خرید این مشتریان نمی توان به قواعد انجمنی این گروه استناد کرد، بنابراین به نتایج خوشه بندی استناد می کنیم. با توجه به نتایج حاصل شده برای مشتریان گروه دوم شرکت کدبانو، شرکت باید ویژگی های منظم و دوره ای را مد نظر قرار داده و با

های بازه زمانی در اختیار، نشان می داد. در مرحله بعد، با ترکیب نتایج حاصل از خوشه بندی کای- میانگین و الگوریتم Apriori به تعیین روابط و الگوی خرید گروه های مختلف پرداخته شد؛ همانطور که مطرح شد، ۱۲ قاعده برای خوشه اول، ۲ قاعده خوشه دوم و سوم، ۷ قاعده خوشه چهارم، ۱ قاعده از خوشه پنج و هفت و ۳ قاعده برای خوشه های شش و هشت استخراج شد. نتایج مجدداً با خبرگان شرکت مورد تجزیه و تحلیل قرار گرفت و سیاست های تخفیف دهی استخراج شدند.

در این تحقیق از یک الگو ترکیبی خوشه بندی- قواعد انجمنی برای ارائه سیاست تخفیف دهی استفاده شد که مزیت آن نسبت به تحقیق های دیگر به شمار می آید. همچنین در تمام مراحل داده کاوی به منظور تجزیه و تحلیل نتایج از حضور خبرگان بازاریابی و فروش شرکت استفاده شده که منجر به قابل فهم و کاربردی بودن نتایج شده است. مطالعه یک شرکت بزرگ پخش مواد غذایی در ایران از دیگر مزایای این تحقیق می باشد که در تحقیق های مشابه داخلی یافت نشد. از محدودیت های تحقیق می توان به دسترسی به داده ها در یک بازه زمانی کوتاه (هفت ماه)، عدم دسترسی به تمام فیلدهای مورد نیاز (به خصوص اطلاعات مالی) اشاره کرد. همچنین به منظور تحقیق های آتی، الگو سازی نتایج حاصل در قالب یک سیستم پشتیبان تصمیم هوشمند پیشنهاد می شود.

منابع

۱. حمیدی، ناصر - بازرگان، ابوالقاسم - جباری، فهیمه " بررسی تاثیر رفتار شهروندی سازمانی کارکنان بر میزان وفاداری مشتریان"، دی ۱۳۹۰، دومانه علمی-تحقیقی دانشگاه شاهد، سال هجدهم، دوره جدید، شماره ۵۰
2. Hanafizadeh Payam, Rastkhiz Paidar Neda, " data mining techniques in customer segmentation : A case study in automobile insurance in an insurance company", Journal of Sanaat - E- Bimeh, vol.26, summer 2011, number2 (102), pp. 55-81.
3. Fridman J. H. (1997) "Data mining and statistics. What's the connection?", Proc. Of the 29th Symposium on the Interface: Computing Science and Statistics, May 1997, Houston, Texas.
4. Han. J, Kamber. M. (2006) "Chapter 1: Introduction", Data mining concepts and techniques, 2nd edition,, Morgan Kaufmann Publishers.

برنامه ریزی و سرمایه گذاری روی این موضوع، تعداد تکرارهای خرید مشتری را افزایش دهد. از طرف دیگر باید برای مشتریانی که دوره ویزیت را رعایت کرده و در هر بار ویزیت اقدام به سفارش دهی می کنند، تخفیف قائل شود. همچنین شرکت کدبانو برای این دو دسته از مشتریان خود (که قبلاً در این تحقیق شناسایی شدند)، دو نوع سیاست گذاری تخفیف روی تعداد اقلام و تخفیف کلی روی خرید را به تفکیک دسته ها اجرا کند چرا که طبق نتایج این تحقیق اعمال این سیاست ها بر این دو دسته منجر به ارتقاء وفاداری مشتریان و افزایش ارزش مالی خرید آنها از شرکت شده که نهایتاً به سودآوری شرکت منجر خواهد شد.

در این تحقیق دو الگو RFM و LRFM به کار گرفته و با هم مقایسه شدند و از ترکیب دو تکنیک داده کاوی، خوشه بندی - الگوریتم کای میانگین- و قواعد انجمنی - الگوریتم Apriori-، جهت دستیابی به رفتار و الگوهای دقیق تر از مشتریان استفاده شد که مزیتی نسبت به تحقیق ها و روش های مشابه می باشد.

نتیجه گیری و پیشنهاد برای تحقیق های آتی

در این تحقیق خوشه بندی مشتریان را با معیار RFM و الگوریتم کای میانگین انجام دادیم؛ و با توجه به نتایج شاخص های ارزیابی دان و SSE، تعداد خوشه بهینه ۸ عدد تعیین شد. با توجه به ارزش گذاری مشتریان بر حسب معیار RFM هشت حالت برای مشتریان ایجاد شد. طی جلساتی که با حضور خبرگان و مدیران بخش فروش و بازاریابی شرکت برگزار شد، نتایج خوشه بندی ها بررسی شد و با توجه به نظر صاحب نظران متغیر تأخر خرید در برچسب گذاری دخیل نشد، با حذف متغیر تأخر خرید، حالات موجود برای مشتریان به چهار گروه مشتریان برتر، مشتریان تکرارشونده، مشتریان خرج کننده و مشتریان نامطمئن کاهش یافت. با توجه به ماتریس ارزش مشتری، خوشه های را به دو گروه مشتریان برتر و نامطمئن برچسب گذاری کردیم. برای پیدا کردن روابط بین کالاها و اینکه کدام کالاها با هم خریداری می شوند، از تکنیک داده کاوی قواعد انجمنی استفاده شد، با تعیین مقدار حداقل اطمینان و پشتیبان در الگوریتم Apriori، ۱۳ قاعده استخراج شد که ارتباط بین کالاها را در کل تراکنش

۱۶. غضنفری، مهدی - ملک محمدی، سمیرا - علیزاده، سمیه - فتح الله، مهدی "بخش بندی مشتریان صادراتی میوه های خوراکی"، فصلنامه تحقیقات بازرگانی، شماره ۵۵، تابستان ۱۳۸۹، ۱۸۱-۱۵۱.
۱۷. بابائی کیایی، مسعود - شهرابی، جمال و سپه وند مهرداد، "بکارگیری تکنیکهای داده کاوی و CLV در رتبه بندی رفتاری مشتریان در صنعت بانکداری مبتنی بر اخذ تصمیمات راهبردک"، ۲۰۱۱؛ پنجمین کنفرانس داده کاوی ایران.
18. Henry Chan, C. C. / Cheng, Chi-Bin / Hsien, Wen-Chen (2011); "Pricing and promotion strategies of an online shop based on customer segmentation and multiple objective decision making"; *Expert System With Applications* p.38.
19. Khajvand, Mahboubeh / Zolfaghar, Kiyana / Ashoori, Sara / Alizadeh, Somayeh (2012); "Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case Study"; *Procedia computer science*, pp. 57-63
20. Der-Chiang Li, Wen-Li Dai, Wan-Ting Tseng (2011). A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business, 38 (2011) pp.7186-7191
۲۱. شمشیری، علیرضا - صباری، سید حسین و تارخ، محمدجعفر "شناسائی و کشف الگوهای خرید مشتریان لوازم خانگی از طریق داده کاوی"، ۱۳۹۰؛ پنجمین کنفرانس داده کاوی ایران.
22. Seyed Hosseini S.M., Maleki Anahita, Gholamian M., "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty", *Expert system with applications*, vol. 37, July 2010, pages 5259-5264.
۲۳. مینائی، بهروز - نصیری، مهدی - حسنی، دانیال - شناسا، ابراهیم "آموزش گام به گام داده کاوی با "Clementine"، گروه مهندسی - تحقیقی ساحر، پارسه، ۱۳۹۰.
24. Agarawal, C. C., & Yu, P. S., (2002). Finding localized associations in market basket data. *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, pp. 51-62.
25. Ahn, Kwang-Il (2012); "Effective product assignment based on association rule mining in retail", *Expert Systems with Applications* p.39.
۵. فرتوت، رقیه - شهرابی، جمال (۱۳۹۰) "ارائه الگوی به منظور شناسائی تأمین کنندگان مناسب با استفاده از داده کاوی"، پنجمین همایش داده کاوی ایران.
6. Tyron, RC. (1939), *Cluster Analysis*, New York: McGrew Hill.
7. Kaufman L, Rousseuw P. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley and sons.
۸. مومنی، منصور - لعلی چشمه، محمدرضا - نالچیگر، سروش "شناسائی نقاط اشباع نشده بازار محصولات ایران خودرو با استفاده از تحلیل خوشه ای و فرآیند تحلیل سلسله مراتبی"، تیر ۱۳۸۹، دومینارمه علمی - پژوهشی دانشگاه شاهد، سال هفدهم، دوره جدید، شماره ۴۱
9. Ahn, Kwang-Il (2012); "Effective product assignment based on association rule mining in retail", *Expert Systems with Applications*. p. 39.
10. Cheng, Ching-Hsue and Chen, You-Shyang (2009); "Classifying the segmentation of customer value via RFM model and RS theory", *Expert systems with applications*, Vol. 36, pp. 4176-4184. Hughes, Arthur M. (1994), 'Strategic database marketing', Chicago: Probus publishing.
11. Cheng, Ching-Hsue and Chen, You-Shyang (2009); "Classifying the segmentation of customer value via RFM model and RS theory", *Expert systems with applications*, Vol. 36, pp. 4176-4184. Hughes, Arthur M. (1994), 'Strategic database marketing', Chicago: Probus publishing.
12. Cheng, Ching-Hsue and Chen, You-Shyang (2009); "Classifying the segmentation of customer value via RFM model and RS theory", *Expert systems with applications*, Vol. 36, pp. 4176-4184. Hughes, Arthur M. (1994), 'Strategic database marketing', Chicago: Probus publishing.
13. Ghorbani Kh., Khalili Ali, Alavipanah S.K., Nakhaeizadeh Gh.R., "Comparative study of the meteorological drought indices (SPI and SIAP) using data mining method (Case studt of Kermanshah Province)", *Journal of water and soil (Agricultural science and technology)*, Aug-Sep 2010, vol. 24, number3, p. 417-426.
۱۴. غضنفری، مهدی - علیزاده، سمیه تیمورپور (۱۳۸۳) داده کاوی و کشف دانش، انتشارات دانشگاه علم و صنعت.
15. Agrawal R., T. Imielinski and A. Swami (1993) "Mining association rules between sets of items in massive databases", *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Washington DC, USA.