

ارائه الگو تحلیل‌های پیشرفته آینده‌نگر براساس ارتباط میان سیاست‌های فروش با خوشه‌بندی مشتریان و سبد محصولی آنها در صنعت فروش و پخش

بابک سهرابی یورتچی^{۱*}، ایمان رئیسی وانانی^۲، یاسمن حاجی قربانی^۳

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

۲- دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

۳- دانشجوی کارشناسی ارشد، گروه مدیریت فناوری اطلاعات، دانشکده مدیریت دانشگاه تهران، ایران.

پذیرش: ۱۳۹۹/۵/۲۵

دریافت: ۱۳۹۹/۲/۲۷

چکیده

راهکارهای آینده‌نگر در سیاست‌گذاری‌های فروش یکی از مؤلفه‌های مهم و ضروری در برنامه‌ریزی‌های کوتاه‌مدت و میان‌مدت در هر کسب‌وکاری است. تدوین درست و دقیق سیاست‌های فروش می‌تواند در مدیریت جریان‌های نقدی و تخصیص منابع نقش مؤثری داشته باشد. امر مزبور حکایت از این دارد که باتوجه‌به اینکه در دنیای رقابتی امروز (که رضایت خاطر مشتریان و افزایش سهم بازار مهم تلقی می‌شود) برآورد مسئله در رابطه با رفتار خرید مشتریان موضوع اصلی هر سازمانی است که با استفاده از ارائه راهکار در جهت تحلیل‌های پیشرفته آینده‌نگر در حوزه سیاست‌گذاری می‌توان این مسائل را بهبود بخشید. هدف این پژوهش ارائه الگوی برای خوشه‌بندی مشتریان و استخراج سبد محصولی هر خوشه است که با استفاده از این اطلاعات سیاست‌های فروش و پخش متناسب هر خوشه استخراج‌شده که در نهایت نتایج به‌دست‌آمده به‌وسیله خبرگان صنعت پخش تأیید



شده است. صرف‌نظر از مشتریان فعلی با ورود مشتریان جدید، با بهره‌گیری از الگوریتم دسته‌بندی خوشه‌ها، سیاست مناسب این دسته از مشتریان نیز پیش‌بینی می‌شود که قابل ارائه برای تقویت و وفاداری آنها خواهد بود.

واژه‌های کلیدی: سیاست‌گذاری فروش، خوشه‌بندی، سبد محصولی، صنعت پخش، داده‌کاوی.

۱- مقدمه

در دنیای رقابتی امروز کشف نیازهای مشتریان، برآورد نیازهای آنها و به‌دنبال آن کسب سود از جمله مهم‌ترین اهداف شرکت‌های حوزه پخش محسوب می‌شود. برای دستیابی به این موضوع مهم باید سیاست‌های توزیع کالا و سیاست‌های فروش محصولات را بهینه‌سازی کرد چرا که تدوین مناسب سیاست‌ها سبب رسیدن به اهداف استراتژیک و رضایت خاطر مشتریان می‌شود که از طرفی نیز می‌توان بهای تمام‌شده و هزینه‌های شرکت و تأمین‌کنندگان را پوشش داد [۱].

امروزه شرکت‌های مطرح در حوزه پخش از راهکارهای مختلف تحلیلی برای حفظ مشتری استفاده می‌کنند و هزینه‌های کلانی برای جذب مشتریان جدید می‌پردازند. بزرگان حوزه بازاریابی اذعان می‌کنند در صورت تمرکز منابع سازمان به جذب مشتریان جدید و تلاش برای نگهداشت آنها می‌توان به مزیت رقابتی در این عرصه دست پیدا کرد. دراین‌راستا کاتلر نیز معتقد است کسب پیروزی در این عرصه زیر سایه فراتررفتن از انتظارات مشتری است. بنابراین در کنار جذب مشتریان جدید، نگهداشت و وفادارساختن مشتریان برای کسب سود بسیار امر کلیدی محسوب می‌شود [۲]. در این پژوهش نخست مشتریان را به‌وسیله الگوریتم کامینز^۱ خوشه‌بندی شده و روی هر خوشه الگوریتم قواعد انجمنی اجرا می‌شود. درنهایت نیز سبد محصولی هر خوشه نیز استخراج می‌شود. در ادامه با کمک انجمن خبرگی و براساس ویژگی موجود در هر خوشه و سبد محصولی آن، سیاست‌های متناسب با هر خوشه تخصیص داده می‌شود که به تقویت فروش آن کمک خواهد کرد. از طرفی با ورود مشتری

1. K-means



جدید با استفاده از الگوریتم دسته‌بندی شبکه عصبی، خوشه‌مناظرش را پیش‌بینی کرده و سیاست مناسب را اتخاذ می‌کنیم که این امر سبب جذب و تقویت این دسته از مشتریان تازه‌ورود می‌شود.

۲- مروری بر ادبیات پژوهش

۲-۱ سیاست‌های پیشبرد فروش^۱

در سال‌های اخیر سیاست‌های فروش و پخش در صنعت پخش بسیار حایز اهمیت است، کالا از تولیدکننده به توزیع‌کننده و از توزیع‌کننده به دست مصرف‌کننده نهایی می‌رسد که در این میان اتخاذ سیاست‌های متناسب در روند فروش بسیار تأثیرگذار خواهد بود. در واقع مجموعه فعالیت‌هایی را که سبب ارتباط مؤثر بین کانال‌های توزیع و انتقال سریع کالا به مصرف‌کننده نهایی و در نهایت فروش خواهد شد، سیاست‌گذاری فروش و پخش می‌نامند [۳].

ستاد فروش در شرکت‌های پخش در اصل علاوه بر حفظ مشتریان فعلی باید به دنبال نفوذ در مشتریان جدید نیز باشند چرا که با افزایش تنوع محصولات و گسترش کانال‌های توزیع و پخش دو موضوع نفوذ و گسترش در بازار امر حیاتی محسوب می‌شود. در این راستا تصمیم و سیاست‌گذاری فروش بر مبنای سبد خرید مشتریان فعلی و پیش‌بینی نیازهای بازار برای جذب مشتریان‌گیری جدید از جمله اقداماتی است که برای رویارویی با فضای رقابتی در این صنعت مؤثر خواهد بود [۴].

۲-۲ داده‌کاوی^۲

در دنیای امروز حجم زیاد داده‌ها و تنوع منابع اطلاعاتی سبب شد که علم آمار محدودیت‌های خود را نشان دهد. در نهایت با افزایش ذخیره‌سازی داده، سرعت زیاد رایانه‌ها و الگوریتم‌های متعدد استفاده‌شده سبب پیدایش علم داده‌کاوی شده است که این علم به استخراج اطلاعات و

1. sales promotion
2. datamining



دانش و کشف الگوی پنهان از یک پایگاه داده‌های بسیار بزرگ می‌پردازد [۵]. امروزه در محیط کسب‌وکار نیز باتوجه به تنوع بازارهای هدف، پیچیدگی الگو کسب وکار و همچنین چالش در دسترسی به داده‌های پراکنده و حجیم سبب شد تا نیاز به استفاده از یک روش علمی و اثبات‌شده که بتواند مدیران را در کشف قوانین پنهان داده‌ها یاری کند، احساس شد. به همین جهت استفاده از روش‌ها و روش‌های داده‌کاوی، این قدرت را به تصمیم‌گیرندگان خواهد داد تا درباره آینده سیستم خود، دانش تحلیلی داشته باشند و برنامه‌های راهبردی را برای آینده شرکت خود تدوین کنند [۶].

چارچوب مفهومی پژوهش جاری متشکل از متغیرهای اثرگذاری است که در بررسی پایگاه داده عملیاتی مشتری شناسایی شده‌اند و شامل فروش ریالی، تعداد فروش، دفعات خرید مشتریان، تعداد و قلم سند فاکتور و سفارش می‌باشد که با استفاده از روش‌های داده‌کاوی الگوسازی انجام می‌شود (بخش‌بندی مشتریان و شناسایی مبانی تفکیک آنها، قواعد انجمنی بر هر خوشه و طبقه‌بندی مشتریان براساس خروجی‌های ایجادشده). درنهایت نیز پس از ارزیابی آنها برای بهبود فرایند سیاست‌گذاری استفاده می‌شود.

۲-۳- دسته‌بندی

در الگوریتم‌های دسته‌بندی که ازجمله روش‌های پیش‌بینی محسوب می‌شود، مجموعه داده اولیه به دو مجموعه داده با عنوان مجموعه داده‌های آموزشی^۱ و مجموعه داده‌های آزمایشی^۲ تقسیم می‌شود.

در مرحله آموزش براساس سایر ویژگی‌ها برای اندازه‌های ویژگی دسته، الگو ساخته می‌شود. پس از ساخت الگو در مرحله ارزیابی، دقت الگو ساخته‌شده به کمک مجموعه داده‌های آزمایشی ارزیابی خواهد شد که ازجمله معیارهای ارزیابی دسته‌بندی دقت در دسته‌بندی، معیار خطای دسته‌بندی و بازخوانی می‌باشد [۷].

یک کسب‌وکار بدون داشتن درکی عمیق از اینکه بهترین مشتریان فعلی شرکت چطور دسته‌بندی شده‌اند، اغلب فاقد تمرکز بازار موردنیاز برای تخصیص مؤثر منابع ارزشمند انسانی

1. train dataset
2. test dataset



و سرمایه خود است. اگر دسته‌بندی بهترین مشتریان فعلی به شکل صحیحی انجام شود، مزایای آن برای کسب‌وکار بی‌شمار است.

از جمله روش‌های برگرفته از یادگیری ماشینی، شبکه عصبی است که ساختار قابل یادگیری دارد، به طوری که خطای خود را اصلاح کند، یادگیری در این سیستم‌ها به صورت تطبیقی می‌باشد. ساختار این الگوریتم عناصر به هم پیوسته دارد که با آسیب دیدن یک قسمت از آن، قسمت دیگر آن را ترمیم کرده و در بازسازی آن مشارکت خواهد کرد. یک شبکه به طور معمول شامل یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی است [۸].

۲-۴- خوشه‌بندی

خوشه‌بندی یکی از انواع روش‌های داده کاوی و از نوع آموزش بدون نظارت است که داده‌هایی که شباهت زیادی دارند در یک خوشه قرار می‌گیرند و هر مورد با گروه خود بیشترین شباهت و با هم‌گروهی‌های خود کمترین شباهت را خواهد داشت [۷]. یکی از روش‌های خوشه‌بندی در داده‌کاوی کامینز است که یکی از پرکارترین روش‌های خوشه‌بندی محسوب می‌شود.

۲-۵- کشف قوانین انجمنی

از جمله روش‌های توصیفی تلقی می‌شود که یکی از کاربردهای مهم آن «تجزیه و تحلیل سبد بازار» می‌باشد. در این راستا خروجی آن کالاهایی است که بیشتر به طور همزمان خریداری می‌شوند. پیشرفت فناوری، فروشگاه‌های خرده‌فروش را قادر کرده است تا با تجزیه و تحلیل سبد بازار خرده‌فروشان، عملکرد خرید مشتریان را پیش‌بینی کنند. با استفاده از این کار مدیران فروش می‌توانند تخفیف‌های بهتری را ارائه دهند و بتوانند بهتر سبد خرید مشتریان خود را مدیریت کنند [۸].



۳- پیشینه پژوهش

به صورت کلی مطالعات انجام شده در این حوزه به صورت خلاصه در جدول ۱ آمده است.

جدول ۱. پیشینه پژوهش

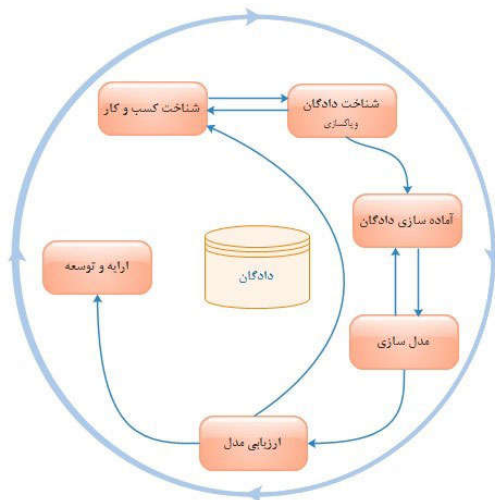
محقق	موضوع	یافته‌های پژوهش
[۲]	ارائه یک الگو ترکیبی داده‌کاوی برای تعیین راهبرد تخفیف‌دهی	<ul style="list-style-type: none"> دستیابی به رفتار و الگوهای دقیق‌تر از مشتریان اعمال سیاست تخفیف به خوشه‌های مشتریان برای ارتقای وفاداری مشتریان و افزایش ارزش مالی خرید آنها
[۹]	بررسی تأثیر بازاریابی رابطه‌مند بر وفاداری مشتریان	<ul style="list-style-type: none"> ارائه معیارهایی برای ایجاد وفاداری در مشتریان و ایجاد مزیت رقابتی
[۱۰]	بهینه‌سازی عملکرد زنجیره تأمین و سیاست‌گذاری‌های تولید محصول برای محصولات جدید	<ul style="list-style-type: none"> افزایش سطح سود، بهینه‌سازی و مدیریت موجودی انبار و تدوین برنامه بازاریابی
[۱۱]	پیش‌بینی فروش محصول با استفاده از بررسی‌های بر خط خرید مشتریان و داده‌های فروش	<ul style="list-style-type: none"> پیش‌بینی فروش محصول با استفاده از بررسی‌های آنلاین و داده‌های فروش تاریخی اطلاعات مفیدی را درخصوص برنامه بازاریابی، مدیریت فروش، لجستیک و انبار در اختیار مدیران قرار می‌دهد.
[۱۲]	پیش‌بینی ارزش دوره حیات مشتری به‌وسیله یک الگو داده‌کاوی	<ul style="list-style-type: none"> کمک و ارتقای CRM در سازمان‌های امروزه و ایجاد مزیت رقابتی، برآورد میزان وفاداری مشتریان و تدوین سیاست‌های فروش بر این مبنا
[۱۳]	پیش‌بینی جذب مشتری با استفاده از درخت تصمیم در «مدیریت ارتباط با مشتری»	<ul style="list-style-type: none"> جلوگیری از اتلاف هزینه‌های سازمان که برای توسعه بازار و کسب سهم بیشتر از بازار قابل کسب است.

۴- روش شناسی پژوهش

در این پژوهش از متدولوژی «کریسپ»^۱ استفاده خواهد شد. این متدولوژی در واقع چارچوب فرایند استاندارد برای انجام طرح‌های داده‌کاوی است.

1. CRISP

چرخه حیات کریسپ از شش مرحله تشکیل شده است. همان‌طور که در شکل ۱ نشان داده شده است، توالی مراحل مستقیم نیست. حرکت به عقب و جلو بین مراحل مختلف همیشه نیاز است [۱۴].



شکل ۱. مراحل متدولوژی چرخه کریسپ

در مرحله فهم تجارت اهداف پروژه و طرح مدیریت پروژه مطرح می‌شود. در مرحله فهم داده کیفیت اطلاعات و توصیف داده‌های پروژه بررسی می‌شود. در مرحله آماده‌سازی داده داده‌ها جمع‌آوری و پاکسازی شده و داده‌های هجو و نامشخص حذف می‌شود. سپس به طراحی ساختار داده پرداخته خواهد شد. در مرحله الگوسازی نیز الگو منتخب اجرا شده ارزیابی می‌شود [۱۵].



۵- اجرای داده‌کاوی

۵-۱- تشریح و آماده‌سازی داده‌ها

۵-۱-۱- انتخاب داده

در این بخش، داده‌ها نخست به صورت سطحی و مقدماتی بررسی شده و مشخصات اولیه مجموعه داده‌ها مانند نوع تعداد متغیرها و تعداد رکورد ها مشخص می‌شود. تعداد ۵۰۰۰ رکورد مربوط به ثبت عملیات فروش است. همچنین این داده‌ها شامل ۱۰ فیلد (سرانه فروش هر مشتری، فروش ریالی و تعداد، سفارش‌های دریافتی ریالی و تعدادی، قلم حواله فروش، تعداد قلم فاکتور، تعداد سند فاکتور و سرانه کل فروش) می‌باشد.

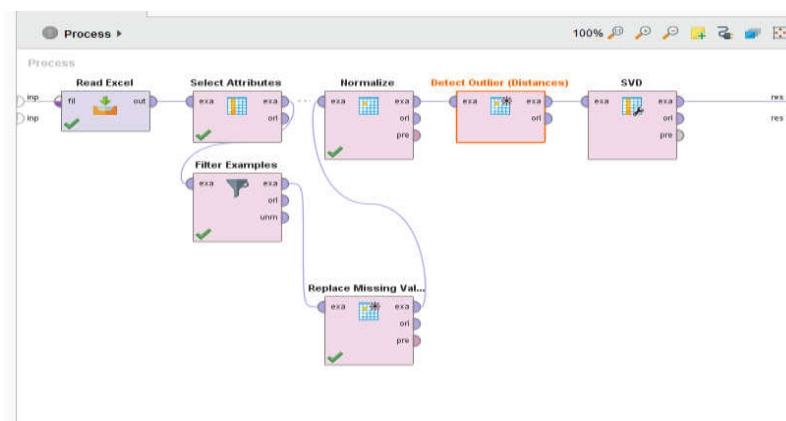
۵-۱-۲- ادغام و ساختار داده

کیفیت داده‌ها کلید کل الگو پیش‌بینی است زیرا می‌تواند بر نتایج پیش‌بینی شده از تجزیه و تحلیل تأثیر بگذارد. از این رو پیش‌پردازش اطلاعات باید قبل از تجزیه و تحلیل داده‌ها به این ترتیب انجام شود (شکل ۲):

- جایگزینی مقادیر ازدست‌رفته با صفر
- دیتا ست شرکت‌های پخش ما دارای تعدادی از مقادیر گم‌شده و مقادیر غیرممکن است، در این مجموعه داده مقادیر ازدست‌رفته با صفر جایگزین شده است.
- جایگزینی مقادیر گم‌شده و نامشخص با میانگین
- در این مطالعه، پیش‌پردازش داده‌ها با جایگزینی مقادیر گم‌شده با میانگین انجام می‌شود. پس از مشخص شدن تعداد آن برای هر رکورد، آنهایی که ۲ و بیشتر از ۲ مشخصه آنها مقداری نداشته باشد، حذف می‌شود.
- نرمال‌سازی داده
- پس از جایگذاری مقادیر نامشخص برای حذف تأثیر وزن ویژگی‌های عددی آنها را در بازه ۰ تا ۱ نرمال‌سازی می‌کنیم.



- شناسایی داده‌های پرت با استفاده از فاصله اقلیدسی با استفاده از فاصله اقلیدسی داده‌های پرت شناسایی می‌شوند. تعداد همسایگی برابر ۲۰۰ و تعداد داده‌های پرت ۱۵۰ در نظر گرفته شده است. به این ترتیب ۱۵۰ رکورد به عنوان داده پرت شناسایی و حذف می‌شوند.
- کاهش ابعاد با استفاده از روش جبر خطی^۱ در این طرح از روش جبر خطی استفاده کرده و تعداد ویژگی‌ها را به ۵ عدد کاهش دادیم.



شکل ۲. پیش پردازش داده‌ها

۶- الگوسازی

۶-۱- خوشه‌بندی و ارزیابی

در این طرح، در هر روش بهینه‌ترین تعداد خوشه را انتخاب کرده و عمل خوشه‌بندی رو انجام دادیم. همان‌طور که در جدول ۲ دیده می‌شود، روش کی مینز به علت آنکه کمترین مقدار را برای شاخص دیویس بولدین داشت، به عنوان بهترین روش انتخاب شد. در روش کی مینز از

1. DSV



عملگر مربوطه^۱ استفاده کرده و بهترین تعداد خوشه را براساس جدول ۲ انتخاب می‌کنیم. در نهایت الگوریتم خوشه‌بندی با تعداد ۶ خوشه برای آن اجرا شد.

جدول ۲. ارزیابی انواع الگوریتم‌های خوشه‌بندی

نام الگوریتم	تعداد خوشه	زمان اجرا	دیویس - بولدین
کامینز	۸	۳۰s	DB : ۰/۲۰۴
کی مدیوز	۶	۱۲M	DB : ۰/۳۴
ایکس مینز	۸	۴۰S	DB : ۰/۴۳۹

۷- نتایج حاصل از خوشه‌بندی مشتریان

نتایج حاصل از خوشه‌بندی نشان براساس جدول ۳ نشان می‌دهد که مشتریان ما به ۶ خوشه با مشخصه‌های عرف شرکت‌های پخش تقسیم می‌شوند.

جدول ۳. خوشه‌بندی مشتریان و مشخصه آنها

خوشه مشتریان	فروش ریالی	دفعه‌های خرید مشتری	فروش تعدادی	تنوع اقلام خرید مشتری
مشتریان با ارزش ریالی بالا	بالا	پایین	بالا	پایین
مشتریان برتر و وفادار	بالا	بالا	بالا	بالا
مشتریان پرتکرار	پایین	بالا	پایین	بالا
مشتریان کم‌تکرار یا نامطمئن	پایین	پایین	بالا	بالا
مشتریان کم‌ارزش	پایین	بالا	پایین	پایین
مشتریان در حال ریزش و یا از دست‌رفته	پایین	پایین	پایین	پایین

1. performance



۸- اجرا و ارزیابی قواعد انجمنی

با استفاده از قواعد انجمنی ترکیب سبد خرید هریک از خوشه‌ها را مطابق جدول ۴ جدا به دست می‌آوریم. به این ترتیب که الگوریتم را برای هریک از ۶ خوشه جدا اجرا می‌کنیم و براساس شاخص‌هایی که این الگوریتم داده است و محصولات موجود در هر سبد خرید و با کمک گروهی از خبرگان و مدیران فروش شرکت‌ها سیاست‌هایی رو برای هر خوشه تخصیص می‌دهیم.

لازم به ذکر است در هر خوشه با کمک دو شاخص پشتیبان و اطمینان سیاست‌ها رو تخصیص می‌دهیم. منظور از شاخص پشتیبان، یعنی کالا چند بار در سبد خرید مشتری تکرار شده و منظور از شاخص اطمینان، یعنی خرید کالا در سبد خرید مشتری چند بار اطمینان داشته است.

جدول ۴. نتایج قواعد انجمنی هر خوشه

شماره خوشه	پشتیبان	اطمینان	تالی	مقدم
خوشه ۱	۰/۲	۰/۹۱	لیوانی ۱۰۰ گرم پاک خامه	شیر کاکائو بطری ۹۴۶ گرم ۳ درصد پاک
	۰/۰۹	۰/۵۱	۲۰۵ گالن ۵۲۰۰ وانیلی بستنی	بستنی ۲۰۵ گالن ۵۲۰۰ سستی خامه‌ای
	۰/۰۹	۰/۵۱	۲۰۵ گالن ۵۲۰۰ وانیلی بستنی	شیر نایلونی ۹۰۰ گرم کم‌چرب پاک
	۰/۰۵	۰/۱	لیوانی ۱۰۰ گرم پاک خامه	پنیر ۱۰۰ گرم گلا آمل
خوشه ۲	۰/۹	۱/۳۳	شیر کاکائو بطری ۹۴۶ گرم ۳٪ پاک	بستنی لیتری ۶۰۰ گرم کیک پسته‌ای، بستنی حصیری ۶۰ گرم وانیل توت فرنگی پاک، ماست لیوانی ۷۵۰ گرم پروبیوتیک ۴/۱ درصد پاک، بستنی زمستانی ظفر
	۰/۹	۱/۳۳	شیر کاکائو بطری ۹۴۶ گرم ۳ درصد پاک	بستنی ۲۰۵ گالن ۵۲۰۰ شکلاتی، ماست لیوانی ۱/۳ درصد ۹۰۰ گرمی پاک، خامه قنادی نایلون ۵۰۰۰ مینارین پاک
	۰/۹	۱/۳۳	شیر کاکائو بطری ۹۴۶ گرم ۳ درصد پاک	بستنی کافی شاپی ۶۰۰۰ گرم شکلات نارگیل، فالوده شیرازی سگال ۹۰۰ پاک، خامه قنادی نایلون ۵۰۰۰ مینارین پاک



شماره خوشه	پشتیبان	اطمینان	تالی	مقدم
خوشه ۳	۰/۲	۰/۱	ماست لیوانی ۹۰۰ سنتی پاک آرا، شیر بطری نیم‌چرب ۲۵۰ گرم پاک، شیر نیم‌چرب ۲۵۰ گرم پاک بطری	بستنی بانجو ۷۰۰ گرم کیک بسته‌ای، کره ۵۰ گرم ۸۲ درصد پاک، بستنی ۲.۵ گالن ۵۲۰۰ توت‌فرنگی، ماست لیوانی ۴۵۰ گرم پروبیوتیک ۴/۱ درصد پاک و خامه لیوانی ۱۰۰ گرم پاک
	۰/۰۴	۰/۵	دوغ ۲۵۰۰ گرم خمره‌ای پاک پی، شیر بطری نیم‌چرب ۲۵۰ گرم پاک	بستنی کافی شاپی ۲۵۰۰ گرم قهوه اسپرسو پاک، کره ۵۰ گرم ۸۲ درصد پاک، خامه لیوانی ۱۰۰ گرم خرما پاک، خامه قنادی شیرین‌شده بانجو پاک پی
	۰/۰۴	۰/۵	بستنی چوبی ۸۵ گرم طلایی پاک، ماست موسیر ۱۰۰ گرم پاک	خامه لیوانی ۱۰۰ گرم پاک، ماست لیوانی ۴۵۰ گرم ۱/۳ درصد پاک، ماست لیوانی ۷۵۰ گرم پروبیوتیک ۴/۱ درصد پاک، ماست لیوانی ۴۵۰ گرم پروبیوتیک ۴/۱ درصد پاک
خوشه ۴	۰/۰۳	۰/۹	خامه لیوانی ۱۰۰ گرمی شکلاتی پاک	خامه لیوانی ۱۰۰ گرمی شکلاتی پاک، کره ۱۰۰ گرم ۸۲ درصد پاک، ماست لیوانی ۷۵۰ گرم پروبیوتیک ۴/۱ درصد پاک، شیر ۱۰۰۰ پت بدون لاکتوز پاک
	۰/۰۴	۰/۱	کره ۵۰ گرم ۸۲ درصد پاک	خامه لیوانی ۱۰۰ گرمی شکلاتی پاک، کره ۱۰۰ گرم ۸۲ درصد پاک، شیر بطری ۲۵۰ گرم پاک، شیر ۱۰۰۰ پت بدون لاکتوز پاک
	۰/۰۴	۰/۱	کره ۵۰ گرم ۸۲ درصد پاک	ماست لیوانی ۷۵۰ گرم پروبیوتیک پاک، شیر بطری ۲۵۰ گرم پاک، شیر ۱۰۰۰ پت بدون لاکتوز پاک
خوشه ۵	۰/۰۲	۰/۲	۱۰۰ گرم گلا امل پنیر	شیر کیسه‌ای ۲۵۰۰ گرم پاک، دوغ پت ۲۵۰ گرم آبعلی (لیبل پاک)، پنیر لیوانی ۲۸۰ لاکتیکی پاک آرا، بستنی ۲.۵ گالن ۵۲۰۰ شکلاتی
	۰/۰۲	۰/۲	بستنی لیوانی ۶۶ گرم وانیلی	شیرکاکائو بطری ۹۴۶ گرم ۳ درصد پاک، ماست لیوانی ۷۵۰ گرم پروبیوتیک ۴/۱ درصد پاک، پنیر لیوانی ۲۸۰ لاکتیکی پاک آرا، خامه استریل اسلیم ۲۰۰ گرم پاکسار



۹- نتایج قواعد انجمنی هر خوشه

گروه اول مشتریانی که فروش ریالی و فروش تعدادی بالایی دارند، اما دفعه‌های خرید و اقلام خرید کمتری دارند و به نام خوشه مشتریان با ارزش ریالی بالا شناخته می‌شوند، برای این گروه باید تخفیف‌های کالاهای مکمل جهت افزایش سبد محصولات آنها ارائه داد. گروه دوم مشتریان برتر و وفادار هستند که فروش ریالی، تعدادی و دفعه‌های خرید بالایی دارند و جزء مشتریان برتر محسوب می‌شوند. برای این گروه باید تخفیفی روی دیگر محصولات قرار داد تا این درصد اطمینان را به ۱۰۰ درصد برسانیم و مشتری کالاهایی را که تمایل دارد، خریداری کند. مشتریان پرتکرار در گروه بعدی قرار دارند. مشتریانی تکرارشونده که دفعه‌های خرید با تنوع بالایی دارند ولی فروش ریالی و تعدادی کمی دارند و در این راستا باید ارزش ریالی خرید آنها را افزایش داد. برای افزایش ارزش مالی خرید این مشتریان باید زمانی که محصولات ماست لیوانی ۹۰۰ سستی پاک آرا، شیر بطری نیم‌چرب ۲۵۰ گرم پاک را خریداری کرد، می‌توان محصول خامه لیوانی ۱۰۰ گرم پاک را به آنها پیشنهاد داد.

خوشه مشتریان کم‌تکرار یا نامطمئن به‌طور عمده مشتریان به‌نسبت جدیدی هستند. برای این دسته از مشتریان که دفعه‌های خرید بسیار پایینی دارند، باید سعی کرد تا ارتباط خود را با خود حفظ کرده و از رقیب و دیگر شرکت‌های پخش دور نگه داریم. در ادامه برای رویارویی با مشتریان مشتریان کم‌ارزش باید تاریخچه خریدهای آنها و سبد خرید آنها را بررسی و ریشه‌یابی کرد. علت پایین بودن ارزش خرید، تعداد خرید پایین و همچنین استفاده نکردن از تنوع محصولات فروشگاه چیست. در نهایت با بررسی سبد خرید این خوشه می‌توان دریافت محصولاتی که در این خوشه هستند، برگشتی زیادی دارند. برای برگشتی‌ها باید بررسی کرد در چه مسیری و به‌وسیله چه ویزیتوری کالا فروش می‌رود. گروه بعدی خوشه مشتریان در حال ریزش و یا ازدست‌رفته است. این گروه، مشتریان خوبی نیستند و نمی‌توان به قواعد انجمنی اتکا کرد.



۱۰- دسته‌بندی مشتریان و ارزیابی الگو

بر اساس دسته‌بندی که از خوشه‌بندی مشتریان و سیاست‌های متناسب آنها داشتیم، با ورود مشتریان جدید (یک دیتاست جدید) همزمان با خریدهای اولیه باید پیش‌بینی کرد که به کدام خوشه تعلق دارند که در این صورت سیاست متناسب با رفتار آنها تخصیص داده می‌شود. مطابق جدول ۵ از میان روش‌های دسته‌بندی روش شبکه عصبی با بالاترین میزان دقت و کمترین خطای طبقه‌بندی قادر به پیش‌بینی اینکه مشتریان جدید مربوط به کدام خوشه می‌باشند، شد.

جدول ۵. ارزیابی روش‌های دسته‌بندی

نام الگوریتم	تکرارپذیری ^۳	کاپا ^۲	دقت ^۱
شبکه عصبی	۸۸/۹۸	۰/۹۹۴	۹۹/۶۰
نایو بیز	۷۲/۳۳	۰/۸۳۷	۸۹/۴۵
نزدیک‌ترین همسایگی	۷۹/۲۵	۰/۹۸۶	۹۸
درخت تصمیم	۰۲/۴۳	۰/۶۱	۵۹/۴۵

۱۱- نتایج حاصل از پیش‌بینی مشتریان جدید

بر اساس نتایجی که از خوشه‌بندی مشتریان حاصل شد، مشتریان به پنج دسته تقسیم و مشتریان جدید به وسیله الگوریتم شبکه عصبی بر اساس جدول ۶ پیش‌بینی شدند. در نتیجه مشتریان جدید هر کدام در خوشه ۲ و ۳ و ۴ و ۵ پیش‌بینی شده و قرار گرفته‌اند و رفتار و سیاست‌گذاری بر اساس خوشه متناظر با آنها انجام خواهد شد که جزئیات آن در بخش‌های قبل ذکر شد.

1. accuracy
2. Kappa
3. recall



جدول ۶. خروجی ارزیابی روش دسته‌بندی

خوشه‌ها	مشتریان با ارزش ریالی بالا	مشتریان برتر و وفادار	مشتریان پرتکرار	مشتریان کم‌تکرار یا نامطمئن	مشتریان کم‌ارزش	مشتریان در حال ریزش و یا از دست‌رفته
تکرارپذیری	۱۰	۰	۴۰	۹۱	۴۷	۸
دقت دسته ^۱	۵	۰	۲۲	۸۸	۳۹	۷

۱۲- محدودیت‌های پژوهش

از جمله مهم‌ترین محدودیت‌ها و چالش‌هایی که در روند اجرای این پژوهش با آنها روبه‌رو بودیم، می‌توان به مسئله دسترسی به داده‌های شرکت‌های حوزه پخش در شرکت بررسی شده به دلیل وجود برخی محدودیت‌ها و مسائل امنیتی، نادر بودن منابع، مقاله‌ها و مطالعاتی که در حوزه داده‌کاوی در سیاست‌های فروش و پخش اشاره کرد. همچنین براساس چرخه حیات گریسپ با روش‌های خوشه‌بندی نمی‌توان تمام نیازهای یک مشکل را به‌صورت همزمان حل کرد و داده‌هایی که ویژگی‌های زیادی دارند، تفسیرهای نتایج آنها یکسان نیست.

۱۳- نتیجه‌گیری

با گذر زمان و افزایش رقبا باید از روش‌های به‌روز و راهکارهای تحلیلی پیشرفته برای جذب بازار استفاده کرد. به این منظور باید باتوجه به گستره زیاد مشتریان و قلم‌های محصولات در صنعت پخش برنامه‌ریزی دقیقی را برای آینده تنظیم کرد تا بتوان در زمان مناسب بهترین واکنش را داشت. برای دستیابی به این امر باید مشتریان را در دسته‌های مختلف جای داد و با علم بر سبد محصولی و اطلاعات در دسترس آنها (فروش تعدادی و فروش ریالی و تنوع محصولات خریداری‌شده و دفعه‌های خرید آنها) سیاست‌های مناسبی را تخصیص داد که به این شکل مشتریان موجود را که وفادار هستند، حفظ کرد و مشتریان ضعیف را تبدیل به مشتریان وفادار کرد و با ورود مشتریان جدید برای آنها سیاست مناسب را تخصیص داد.

1. class precision



در این پژوهش به خوشه‌بندی مشتریان دیتابیس شرکت منتخب حوزه پخش پرداخته شد. در هر روش بهینه‌ترین تعداد خوشه انتخاب و عمل خوشه‌بندی نیز انجام شد. روش کی‌میز به علت آنکه کمترین مقدار را برای شاخص دیویس بولدین داشت، به‌عنوان بهترین روش انتخاب شد. با استفاده از قواعد انجمنی ترکیب سبد خرید هر یک از خوشه‌ها به‌دست آورده شد و براساس شاخص‌های این الگوریتم و محصولات موجود در هر سبد خرید و با کمک گروهی از خبرگان و مدیران فروش شرکت‌ها سیاست‌هایی رو برای هر خوشه تخصیص دادیم. در ادامه با ورود مشتریان جدید به‌وسیله الگوریتم دسته‌بندی نخست خوشه مشتریان جدید مشخص شد. در ادامه نیز سیاست فروش مربوط به خوشه انتخابی برای مشتریان جدید با هدف افزایش مشتریان وفادار به کار گرفته شد. به‌این‌ترتیب علاوه بر حفظ مشتریان فعلی، مشتریان جدید نیز جذب شدند که درنهایت در بازار رقابتی سهم بیشتری کسب خواهد شد.

۱۴- منابع

- [۱] مشهدی حاجی علی ف.، رفیعی م.، علیمیری م. «الگوی تصمیم‌گیری شرکت‌های کوچک و متوسط دانش‌بنیان در شرایط پیچیده و عدم قطعیت عمیق»، مدیریت نوآوری، ۲(۲)، ۱۳۹۲-۱۳۱-۱۲۲.
- [۲] خدیور آمنه، حامدی پروشات «ارائه الگو ترکیبی داده‌کاوی با استفاده از قواعد انجمنی و خوشه‌بندی برای تعیین استراتژی تخفیف‌دهی»، مطالعه موردی شرکت پگاه، صص، ۱۳۹۵، ۱۰-۵.
- [۳] حسینی سرخوش سیدمهدی، اخوان پیمان، دهقانی‌پوده حسین «افزایش موفقیت توسعه محصول جدید مبتنی بر رویکرد نوآوری باز»، نشریه علمی - پژوهشی سال دوم شماره ۲، مدیریت نوآوری، ۱۳۹۲، ۳۰۰-۳۵۰.
- [4] Byrne, Teresa M. McCarthy, Mark A. Moon, and John T. Mentzer. "Motivating the industrial sales force in the sales forecasting process." *Industrial Marketing Management* 40, 2011, 128-138.



[5] خواجوی شکراله، ابراهیمی مهرداد «الگوسازی متغیرهای اثرگذار بر کشف تقلب در صورت‌های مالی با استفاده از تکنیک‌های داده‌کاوی»، فصلنامه حسابداری مالی، ۹(۳۳)، ۱۳۹۶، ۲۳-۵۰.

[6] Chiang, Wen-Yu. "Applying data mining for online CRM marketing strategy." *British Food Journal*, 2018, 45-66.

[7] Majumdar, Jharna, Sneha Naraseeyappa, and Shilpa Ankalaki. "Analysis of agriculture data using data mining techniques: application of big data." *Journal of Big data* 4, no. 1, 2017, 20-45.

[8] Alcalá, Rafael, María José Gacto, and Jesús Alcalá-Fdez. "Evolutionary data mining and applications: A revision on the most cited papers from the last 10 years" *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, no. 2, 2018, 123-190.

[9] قاضی‌زاده مصطفی، بیگی‌نیا عبدالرضا، عباسپور مهنوش، عنبری احسان «بررسی تأثیر بازاریابی رابطه‌مند بر وفاداری مشتریان با استفاده از الگو توسعه‌یافته بازاریابی رابطه‌مند (ERM) (مورد مطالعه: شرکت بیمه ملت)»، *راهبردهای بازرگانی*، ۹(۴۷)، ۱۳۹۰، ۴۰۷-۴۲۲.

[۱۰] احمدی‌دهرشید کیوان، عبدالله‌زاده مقدم سهراب «الگو ریاضی یکپارچه برای بهینه‌سازی سیستم تولید- توزیع زنجیره تأمین کالاهای فاسدشدنی با انبارهای میانی»، *مدیریت تولید و عملیات*، ۱۳۹۸، ۳۷-۵۳.

[11] Tan Y., Li X., Chen Z. "Application of mining association rules in web information retrieval system", *Computer Engineering*, 2006, 57-61.

[12] Linoff G. S., Berry M. J. *Data mining techniques: for marketing, sales, and customer relationship management*, John Wiley & Sons, 2011, 45-90.

[13] Adamson B., Dixon M., Toman N. "The End of Solution Sales.", *Harvard Business Review*, 2012, 310-311.

[14] Shaniqua U., Qaiser H. "A comparative study of data mining process models (KDD, CRISP-DM and SEMMA)", *International Journal of Innovation and Scientific Research*, 12(1), 2014, 217-225

[15] Reischl R. M. "Data mining tools. Data Mining and Knowledge Discovery", *International Journal of Information Technology & Decision Making*, 7(04), 2011, 639-682.