دوره ۱۴، شماره بهار و تابستان ۱۴۰۰

مجله فناوری اطلاعات در طراحی مهندسی
Information Technology in Engineering Design
http://ited.sinaweb.net

تحلیل سبد خرید مشتریان فروشگاه‌های آنلاین با استفاده از قواعد وابستگی و الگوریتم Apriori

رضا مولایی فرد *(۱)

(۱) گروه مهندسی کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران*

(تاریخ ارسال : ۱۴۰۰/۰۱/۱۴ تاریخ پذیرش : ۱۴۰۰/۰۶/۲۹)

چکیده

رشد روزافزون فناوری اطلاعات تحول وسیعی را در فروش کالا و خدمات ایجاد نموده است. با اطلاعات به‌دست‌آمده از مشتریان فروشگاه‌های زنجیره ایی می‌توان از این اطلاعات هم برای بهبود خریدهای آتی این مشتریان هم برای کمک به فروشگاه‌های بزرگ جهت رفع نیاز مشتریان استفاده نمود. برای انجام این کار از روش‌های داده‌کاوی استفاده می‌گردد. تحلیل سبد خرید مشتریان یکی از مهم‌ترین کاربردهای داده کاوی می‌باشد که بر کشف الگوهای خرید به وسیله داده‌های معاملات انجام‌شده، تمرکز می‌کند. در این تحقیق به ارائه روشی به‌منظور بهبود تحلیل سبد خرید مشتریان فروشگاه‌های زنجیره ایی پرداخته می‌شود که تا حد زیادی می‌تواند تحلیل سبد خرید مشتریان را بهبود بخشد. روش پیشنهادی موجود در این تحقیق استفاده از دو الگوریتم قواعد وابستگی و الگوریتم Apriori می‌باشد. طبق تحقیقات انجام‌شده روش پیشنهادی موجود در این تحقیق می‌تواند تا حدود ۹۴٪ به‌درستی سبد خرید مشتریان را تحلیل کرده و نیازهای مشتریان و فروشگاه‌ها را به‌درستی تشخیص دهد.

کلمات کلیدی : تحلیل سبد خرید، فروشگاه زنجیره ایی، داده کاوی، الگوریتم قواعد وابستگی، الگوریتم Apriori

* **عهده‌دار مکاتبات:**

رضا مولایی فرد

نشانی: گروه مهندسی کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران

پست الکترونیکی: rezamolae4@gmail.com **تلفن : ۰۹۳۸۸۴۱۲۰۸۳**

۱- مقدمه

امروزه در دنیای فناوری حجم بسیار زیادی از اطلاعات خام وجود دارد که به‌تنهایی هیچ کاربردی ندارند، اما می‌توان با استفاده از تکنیک‌های داده‌کاوی بهترین بهره‌برداری را از این اطلاعات خام به عمل آورد. داده‌کاوی یک ابزار مطلوب است که می‌توان با استفاده از آن بهترین الگوها و اطلاعات را از داده‌های خام استخراج کرد. با افزایش رقابت بین شرکت‌ها و فروشگاه‌های بزرگ برای حفظ بقای خود با استفاده از داده‌های خام و تکنولوژی اطلاعات به‌پیش بینی وضعیت بازار طی چند ماه آینده بپردازند. این فرایند منجر می‌شود که شرکت‌ها تصمیمات مهمی را در محیط کاری خود اخذ کنند که در پیشرفت آن‌ها بسیار مؤثر است. حضور در بازار پرتلاطم رقابت، سازمان‌ها را با استفاده از رویکردهای نوین جهت نگه داشت مشتریان جاری، شناسایی و جذب مشتریان جدید سوق داده است. روش‌های داده‌کاوی که قابلیت کشف دانش از انبوه داده‌ها را فراهم می‌کنند. در این موضوع به کمک سازمان‌ها و فروشگاه‌ها آمده و آن‌ها در راستای رسیدن به اهداف بازاریابی و مدیریت ارتباط مشتری شامل شناسایی، جذب، نگهداشت و ارتقای مشتری یاری می‌رساند. تحلیل سبد خرید یکی از مهم‌ترین کاربردهای داده‌کاوی است که بر کشف الگوهای خرید به‌وسیله داده‌های معاملات انجام‌شده، تمرکز می‌کند. مدیریت ارتباط با مشتری کارامد تنها در صورتی حاصل می‌شود که سازمان‌ها و فروشگاه‌ها درک درستی از ترجیحات و نیازمندی‌های مشتری داشته باشد. این دانش می‌تواند منجر به تدوین استراتژی سودآور برای فروشگاه‌ها گردد؛ بنابراین فروشگاه‌های به دنبال راهی برای شناسایی هرچه بهتر مشتری و درک نیازهای وی می‌باشند. بدین منظور، ابزار قوی فراتر از بازیابی ساده اطلاعات، جهت تحلیل و بررسی موقعیت مشتری موردنیاز است. در بسیاری از فروشگاه‌ها تبادلات، تنها منبع اطلاعات فروش هستند که داده‌کاوی می‌تواند از آن استفاده کند. داده‌کاوی در کاربردهای متنوعی مورداستفاده قرار می‌گیرد و یکی از ابزارهای مدیریت ارتباط با مشتری می‌باشد. اهمیت به‌کارگیری تکنیک‌های داده‌کاوی در استخراج دانش از حجم عظیم داده‌های فروش و تأثیر استفاده از این اطلاعات در سودآوری و بهبود سیستم مدیریت ارتباط با مشتری است. انتخاب سبد بازار یک فرایند تصمیم‌گیری است که مشتری اقلامی را از میان تعدادی کالا در فرایند خرید انتخاب می‌کند. وابستگی‌های درونی در ارتباط تقاضا در میان اقلام، مشخصه ایی کلیدی در انتخاب سبد بازار است. در این تحقیق به ارائه روش جدیدی به‌منظور تحلیل سبد مشتریان فروشگاه‌های زنجیره ایی پرداخته می‌شود که می‌تواند تا حدود زیادی مشکلات روش‌های پیشین که عدم شناسایی دقیق نیاز مشتریان فروشگاه‌های زنجیره ایی است را برطرف کرده و تحلیل دقیق‌تری را جهت شناسایی نیاز مشتری و هم کمک به ارائه دقیق‌تر و پیشنهاد این محصولات به مشتریان انجام دهد. در این مقاله با به‌کارگیری روش‌های داده‌کاوی در تحلیل سبد بازار با استفاده از قواعد وابستگی و الگوریتم Apriori، به تعیین سبد بازار در فروشگاه‌های زنجیره ایی با جامعه بزرگ از مشتریان پرداخته‌شده است. نتیجه این فرایند داده‌کاوی ارائه مدلی جهت استخراج روابطی به‌منظور تشخیص سبد خرید مشتریان، جهت بهبود عملکرد، افزایش رضایتمندی مشتری و سودآوری در این فروشگاه‌ها خواهد بود

۲- کارهای پیشین

هانگ و همکاران خود در سال ۲۰۲۱ به بررسی تجزیه و تحلیل سبد خرید مشتریان و سپس پیشنهاد این محصولات به سایر مشتریان پرداختند. این محققان معتقد بودند با توجه به مزایای برنامه‌های کوچک که نیازی به بارگیری و نصب ندارند و منابع کمتری را اشغال می‌کنند و اجرای

آن‌ها راحت است، استفاده از برنامه‌های کوچک برای توسعه خرید مشتریان می‌تواند نیازهای خرید کاربران را بهتر برآورده کند. سبد خرید یک ماژول مهم در یک مرکز خرید است. به‌عنوان‌مثال در برنامه WECHAT می‌توان از حافظه پنهان داده و پایگاه داده ابری برای ذخیره داده‌های سبد خرید استفاده کرد و از این اطلاعات برای پیشنهاد به سایر مشتریان استفاده نمود[۱].

ترحینی و همکاران در مقاله خود در سال ۲۰۱۸ به شناسایی عوامل خرید و تحلیل سبد خرید مشتریان پرداختند. هدف این محققان شناسایی عوامل مؤثر بر تأثیر پذیرفتن خرید آنلاین در انگلیس از طریق تلفیق نظریه یکپارچه پذیرش و استفاده از فناوری و مدل DeLone-McLean موفقیت با اعتماد، تنوع محصول و ضمانت محصول داده‌ها از ۳۸۸ مصرف‌کننده خرید آنلاین انگلیسی با استفاده از پرسشنامه خود مدیریت ساختاریافته می‌باشد. مدل‌سازی معادلات ساختاری نشان داد که قصد رفتاری تحت تأثیر امید به عملکرد، راحتی، کیفیت خدمات، اعتماد، تضمین محصول، تضمین محصول و سازگاری، به ترتیب تأثیرشان بر قدرت و درمجموع ۷۰.۴٪ از واریانس BI را تشکیل می‌دهد و تا حدود زیادی می‌تواند سبد خرید مشتریان را به‌درستی تجزیه‌وتحلیل کند[۲].

هانینن و اهلبوم در مقاله خود در سال ۲۰۱۹ به بررسی روش‌های کارآمد برای تجزیه‌وتحلیل سبد خرید مشتریان پرداختند. این محققان آزمایشی را برای درک اثربخشی قیمت‌گذاری پویا بر قصد خرید پس‌ازاینکه کالایی قبلاً در یک سبد خرید مجازی رها شده بود و بررسی تعدیل‌کنندگان این اثر را انجام دادند. نتایج حاصل از این تحقیق مشخص می‌کند که تغییر در قیمت کالاهای باقیمانده در یک سبد خرید مجازی منجر به افزایش قصد خرید در صورت کاهش قیمت می‌شود، که با درک اقتصاد خرد از قیمت‌گذاری و کشش تقاضا مطابقت دارد. از سوی دیگر، همچنین متوجه می‌شویم که انواع مختلف ارتباطات قیمت‌گذاری (کلامی در مقابل تصویری) در افزایش قیمت خریدار بر قصد خرید مصرف‌کننده تأثیر می‌گذارد و درواقع تعداد قابل‌توجهی از مصرف‌کنندگان علی‌رغم قیمت اعلام‌شده مایل به خرید محصول هستند، پس از مراجعه قبلی آن‌ها افزایش یابد. نتایج این تحقیق حاکی از قیمت‌گذاری در یک فضای خرده‌فروشی آنلاین و نقش ریسک‌پذیری و تصمیم‌گیری منطقی مصرف‌کننده به‌عنوان ناظر برای قصد خرید پس از تغییر قیمت اعلام‌شده کمک می‌کند [۳].

اسلامی و همکاران خود در سال ۲۰۲۱ به بررسی تجزیه‌وتحلیل سبد خرید مشتریان بر فروشگاه‌های آنلاین پرداختند. این محققان هدف از این مطالعه را ارائه یک مدل مفهومی یکپارچه با زیربنای مدل پذیرش فناوری گذاشتند تا نشان دهد چگونه زمینه تجربی تعامل وب‌سایت فروشگاه‌های خرده‌فروشی آنلاین بر پاسخ مشتری به خرده‌فروشان آنلاین تأثیر می‌گذارد. این مطالعه با استفاده از آزمایش‌های آزمایشگاهی (آزمایش‌های میدانی نیمه آنلاین)، تأثیرات تعامل وب‌سایت بر درک مصرف‌کنندگان آنلاین از سایت‌های خرده‌فروشی آنلاین را بررسی می‌کند. این طرح آزمایشی موضوعی در بین ۲۹۵ دانشجو از یک دانشگاه بزرگ در استرالیا انجام شد. نتایج نشان داد تعامل وب‌سایت باعث افزایش درک مشتری از سودمندی و سهولت استفاده از وب‌سایت‌های خرده‌فروشی می‌شود. سودمندی درک شده یک وب‌سایت قصد خرید را افزایش می‌دهد و با تأثیر مشتری در خرید آنلاین این تأثیر را تعدیل می‌کند. جالب‌توجه است، نتایج آزمون اقتضایی نشان می‌دهد که تأثیرات مشارکت در خرید آنلاین بر روی پیوند قصد خرید PU در سطح بالایی از تعامل وب‌سایت زیاد است [۴].

وو و یو در مقاله خود در سال ۲۰۲۰ به بررسی رفتارهای مشتریان با انواع مختلف نیازهای اقتصادی، صفحات وب مربوط به RF و ویژگی‌های غیر RF را با استفاده از خوشه‌بندی تجزیه‌وتحلیل پرداختند تا نتایج ارزیابی الگوهای جستجو را با رفتارهای عبور صفحه مرتبط کنند. این تحقیق چهار گروه از مصرف‌کنندگان را در بر می‌گیرد که به جستجوی اطلاعات می‌پردازند. نتایج نشان می‌دهد که مصرف‌کنندگان با نیازهای هدف‌گرا قوی دارای ساده‌ترین مسیرهای جستجو در مقایسه با سایر گروه‌ها هستند، درحالی‌که مصرف‌کنندگان مبتنی بر اکتشاف پیچیده‌ترین مسیرهای

جستجو را دارند. علاوه بر این، مصرف کنندگان با نیازهای بالاتر تمایل به جستجوی مستقیم دارند. در حالی که مصرف کنندگان با نیازهای مبتنی بر اکتشاف تمایل دارند دسته بندی محصولات را کشف کنند و سلسله مراتب طبقه بندی محصولات را به‌عنوان محور بررسی ویژگی‌های وب و سپس انواع خاصی از RF را اتخاذ کنند. نتایج نشان می‌دهد که هر گروه ویژگی‌های وب خاص خود را برای تسهیل روند خرید به کار می‌گیرد و می‌توان انواع مصرف‌کننده را بر اساس رفتار خرید در مرحله اولیه خرید شناسایی کند. این نشان می‌دهد که فروشندگان فروشگاه‌های الکترونیکی می‌توانند ویژگی‌های وب را اصلاح‌کرده و استراتژی‌های بازاریابی متناسب با الگوهای جستجو را برای سطوح مختلف به‌کارگیرند[۵].

۳- ادبیات تحقیق

۳-۱- شناخت داده‌ها

در این مرحله مواردی از قبیل موجود بودن داده‌ها، محرمانگی داده‌ها، چگونگی دسترسی به داده‌ها و میزان سیستمی بودن داده‌ها بررسی می‌شود. تحلیل صحیح در این مرحله ممکن است به تجدیدنظر در مرحله شناخت کسب‌وکار بیانجامد [۶،۷]. جمع‌آوری و آشنایی با داده‌ها، شناخت مسائل مربوط به کیفیت داده‌ها و مشاهده اطلاعات اولیه نهایی یا زیرمجموعه‌های داده‌ها که ممکن است برای آنالیز جالب باشند از حمله این مراحل است. گام‌های این مرحله عبارت‌اند از:

- جمع‌آوری اولیه داده‌ها
- توصیف داده‌ها
- شناسایی داده‌ها
- تائید کیفیت داده‌ها

۳-۲- مدلسازی

این مرحله پیاده‌سازی تکنیک‌های مدلسازی یا داده‌کاوی بر روی دیدگاه قابل کاوش ایجادشده است. گام‌های این مرحله عبارت‌اند از:

- انتخاب تکنیک‌های مدلسازی (فرضیه‌ها و تکنیک های مدلسازی)
- ارزیابی طراحی (طراحی آزمایش)
- ساخت مدل (پارامترهای انتخاب شده و توصیف مدل)
- ارزیابی مدل (سنجش مدل و بازنگری در پارامترهای انتخاب شده)

۳-۳- تبدیل بازدیدکنندگان به خریدار

سیستم‌های توصیه گر در فروشگاه‌های وب‌سایت‌ها با ارائه پیشنهاد کالاهای جدید مرتبط با سبد خرید مشتریان باعث فروش کالاهای بیشتری می‌شوند. درصورتی‌که این پیشنهادات صحیح و مناسب باشند می‌توانند منجر به افزایش چشمگیر فروش محصولات در فروشگاه وب‌سایت شوند.

۳-۴- وفاداری

در دنیای وب و کسب‌وکارهای الکترونیکی فاصله هر فروشگاه با فروشگاه دیگر به‌اندازه یک یا دو کلیک است و این بدین مفهوم است که درصورتی‌که مشتری با صرف اندکی وقت کالای موردنظر خود را نتواند پیدا کند به‌آسانی به فروشگاه دیگری مراجعه می‌کند. سیستم‌های توصیه گر با ایجاد یک رابطه ارزش‌افزوده بین مشتری و وب‌سایت‌ها باعث افزایش وفاداری مشتریان می‌شود. وب‌سایت‌ها بر روی درک نیازهای مشتریان خود و یادگیری سلایق آن‌ها از طریق سفارشی‌سازی‌ها و شخصی‌سازی‌ها سرمایه‌گذاری می‌کنند و مشتریان نیز با برگشت به سایت‌هایی که نیاز آن‌ها را به‌درستی شناخته‌اند به این سرمایه‌گذاری پاسخ می‌دهند. درواقع ایجاد یک رابطه بین مشتریان و سایت باعث افزایش وفاداری آن‌ها می‌گردد و این وفاداری باعث ایجاد یک مزیت رقابتی می‌شود [۸،۹،۱۰].

۳-۵- جذب مشتری

پس از شناسایی بخش مشتریان بالقوه، سازمان‌ها می‌توانند تلاش‌ها و منابع را در راستای جذب مشتریان هدف هدایت کنند. یکی از عناصر جذب مشتریان بازاریابی مستقیم است. بازاریابی مستقیم به ترغیب مشتریان از طریق مختلف جهت سفارش می‌پردازد. نامه الکترونیکی مستقیم و توزیع و تخفیف از انواع بازاریابی مستقیم به شمار می‌رود [۲۸].

۳-۶- نگهداشت مشتری

می‌توان این مرحله را دغدغه اصلی مدیریت ارتباط با مشتری دانست. برآورده کردن انتظارات مشتری در راستای کسب رضایت وی از شروط اصلی نگهداشت مشتری به شمار می‌رود. عناصر اصلی این بخش بازاریابی یک‌به‌یک، برنامه‌های وفاداری و مدیریت شکایات را شامل می‌شود. بازاریابی یک‌به‌یک به دنبال کشف تغییرات در رفتار مشتری و تغییر استراتژی‌ها، مبتنی بر این رفتار است. برنامه‌های وفاداری، پشتیبانی مشتری را به گونه ایی که منجر به یک رابطه دراز مدت با مشتری گردد، در بر می‌گیرد [۲۸].

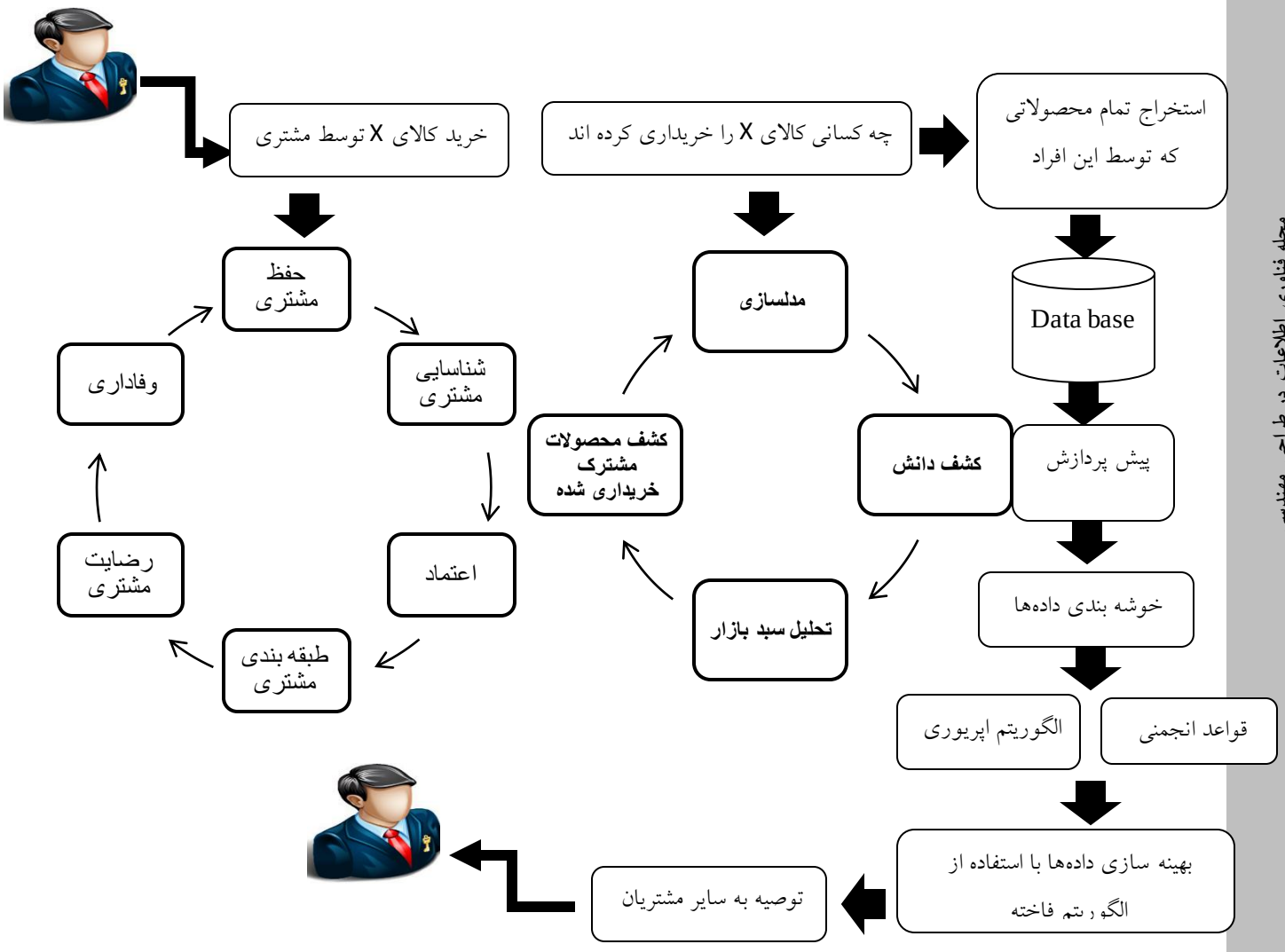
۳-۷- ارتقای مشتری

این مرحله به توسعه روابط با مشتریان شامل بالا بردن ارزش تراکنش‌ها و سودآوری مشتری می‌پردازد. این وظایف با تحلیل ارزش طول عمر مشتری، فروش جانبی و تحلیل سبد خرید مشتری محقق می‌شود. تحلیل ارزش طول عمر مشتری، پیش‌بینی درآمدی است که از مشتری انتظار می‌رود. تحلیل سبد خرید مشتری به دنبال الگوهای متداول در سبد خرید مشتریان است تا تعامل و ارزش مشتریان را حداکثر سازد [۲۸].

۴- روش پیشنهادی

درروش پیشنهادی به ارائه روش جدیدی به‌منظور تجزیه‌وتحلیل سبد خرید مشتریان فروشگاه‌های آنلاین به استفاده از روش‌های داده‌کاوی پرداخته می‌شود. روش پیشنهادی بدین‌صورت می‌باشد که ابتدا پایگاه داده ایی از خریدهای کاربران را که جمع‌آوری می‌کنیم. سپس به‌صورت رندوم خرید گروهی از مشتریان را مورد آنالیز قرار می‌دهیم و محصولات خریداری‌شده این مشتریان رو موردبررسی قرار می‌دهیم سپس به

جستجوی این موضوع که موارد خریداری شده این لیست توسط کدام مشتریان مورد خریداری قرارگرفته است. سپس در مرحله بعد به استخراج تمام محصولات که توسط این افراد خریداری شده اند می‌پردازیم. در مرحله بعد اطلاعات بدست آمده از مراحل قبل را مورد پیش پردازش قرار می‌دهیم. سپس در مرحله بعد داده‌های خود را خوشه‌بندی می‌کنیم تا به گروهای مشابهی از خردی مشتریان دست پیدا کنیم. در مرحله آخر با استفاده از دو روش قواعد وابستگی و الگوریتم Apriori تجزیه‌وتحلیل نهایی محصولات موردبررسی قرار می‌گیرد و درنهایت لیست نهایی و بیشتر مورداستفاده گرفته‌شده موجود در بانک اطلاعاتی موردنظر استخراج‌شده و از این اطلاعات می‌توان برای توصیه به سایر مشتریان استفاده کرد.



شکل ۱. نمایی از روش پیشنهادی

۴-۱- داده‌ها

جامعه آماری این پژوهش یک پایگاه داده از خریدهای مربوط به ۳ ماهه مشتریان فروشگاه‌های بزرگ در سطح کشور می‌باشد. در نتیجه جامعه آماری این تحقیق یک جامعه نامحدود است که به روش تصادفی تعداد ۷۵۰۰ مشتری به‌صورت رندوم انتخاب گردیده‌اند. توصیف کالا به‌عنوان ورودی الگوریتم در نظر گرفته‌شده، این ویژگی‌ها شامل نام کالا، عنوان تجاری کالا و سایر اطلاعات توصیفی می‌باشد که به‌عنوان یک عنصر برای اعمال الگوریتم در نظر گرفته می‌شود و نتیجه الگوریتم به‌صورت قوانینی به شکل X و Y خواهد بود که بیانگر این است که خرید کالای X خرید کالای Y را نتیجه می‌دهد که با توجه به پایگاه داده موجود به پیاده‌سازی روش پیشنهادی می‌پردازیم.

۴-۲- آماده‌سازی داده‌ها

به‌منظور آماده‌سازی داده‌ها برای استفاده در فرایند داده‌کاوی لازم است تا داده‌های خام موجود در فاکتورها را ساختاربندی کنیم و بعد از اتمام پروسه آماده‌سازی داده‌ها، آن‌ها را به شکل مطلوب برای واردکردن به الگوریتم‌های داده‌کاوی آماده کنیم. برای رسیدن به جواب‌های بهینه از حجم داده‌های بسیار بالایی استفاده‌شده است که باید این داده‌ها را مورد پیش‌پردازش قرار دهیم [۱۱].

۴-۳- پیش‌پردازش داده‌ها

در مرحله اول روش پیشنهادی ابتدا باید عمل پیش‌پردازش داده‌ها را انجام دهیم. اگر داده‌های مختلف پیش‌پردازش شوند، عملکرد قابل‌اعتماد و مؤثر یکسانی در تمام مجموعه‌های داده ایی به وجود خواهد آمد [۱۲]. در فرایند داده‌کاوی مانند طبقه‌بندی و خوشه‌بندی نیاز داریم تا داده‌ها برای الگوریتم آماده شوند، زیرا معمولاً نمی‌توان داده‌ها را به‌صورت خام به الگوریتم‌های داده‌کاوی تزریق کرد [۱۳]. برای پیش پردازش داده‌ها، نیاز است تا آن‌ها را از شکل و حالت اولیه، خارج کرده و به شکلی که برای الگوریتم مناسب باشد تبدیل کرد.

۴-۴- کاوش داده‌ها

پس از آن‌که مجموعه داده‌ها از مرحله پیش‌پردازش عبور کردند و برای کاوش آماده شدند بخش دوم فرایند که همان کاوش داده‌ها و کشف الگوهای موجود در این داده‌هاست آغاز می‌گردد. برای کشف این الگوها می‌توان از خوشه‌بندی و ابزارهای داده‌کاوی استفاده نمود که در مراحل بعد جهت کشف الگوها از این الگوریتم‌ها استفاده خواهیم نمود.

۴-۵- خوشه‌بندی داده‌ها

در مرحله بعد باید داده‌های پیش‌پردازش شده از مرحله قبل را خوشه‌بندی کنیم. روش خوشه‌بندی مورداستفاده در تحقیق استفاده از یک روش خوشه‌بندی ترکیبی می‌باشد. بدین‌صورت که ابتدا داده‌ها را با استفاده از الگوریتم خوشه‌بندی C-means، خوشه‌بندی می‌کنیم. سپس با استفاده از خوشه بندی K-means یک خوشه‌بندی مجدد روی داده‌ها انجام می‌دهیم که باعث افزایش دقت خوشه‌بندی می‌شود.

۴-۵-۱- خوشه‌بندی با c-means

در این الگوریتم تعداد خوشه‌ها (c) از قبل مشخص‌شده است. تابع هدفی که برای این الگوریتم تعریف‌شده به‌صورت زیر می‌باشد.

$$J = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{jk}^{m}\, d_{ik}^{2} = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{jk}^{m} \|x_k - v_i\|^2$$ رابطه(۱)

در فرمول فوق m یک عدد حقیقی بزرگتر از ۱ است که اکثر موارد برای m عدد ۲ انتخاب می‌شود. X_k ه Kام است و V_i نماینده یا مرکز خوشه iام است. U_{ik} میزان تعلق نمونه iام در خوشه kام را نشان می‌دهد. علامت $\|*\|$ میزان تشابه (فاصله) نمونه از مرکز خوشه می‌باشد که می‌توان از هر تابعی که بیانگر تشابه نمونه و مرکز خوشه باشد را استفاده کرد. از روی U_{ik} می‌توان یک ماتریس U تعریف کرد که دارای Cسطر و n ستون می‌باشد و مؤلفه‌های آن هر مقداری بین ۰ و ۱ را اختیار کنند اما مجموع مؤلفه‌های هر یک از ستون‌ها برابر ۱ باشد و داریم [۱۴،۱۵،۱۶]. :

$$\sum_{i=1}^{c} u_{ik} = 1, \forall k = 1, \dots, n$$ رابطه (۲)

معنای این شرط این است که مجموع تعلق هر نمونه به c خوشه باید برابر ۱ باشد. با استفاده از شرط فوق و مینیمم کردن تابع هدف خواهیم داشت:

$$u_{ik} = \frac{1}{\sum_{k=1}^{n}\left(\frac{d_{ik}}{d_{jk}}\right)2/^{(m-1)}}$$ رابطه (۳)

$$\boldsymbol{v_i} = \frac{\sum_{k=1}^{n} \boldsymbol{u_{ik}^{m}}}{\sum_{k=1}^{n} \boldsymbol{u_{ik}^{m}}}\ \boldsymbol{x_k}$$ رابطه (۴)

۴.۵.۲. خوشه بندی مجدد با K-means

در مرحله بعد خوشه‌های به‌دست‌آمده از مرحله قبل را به‌صورت مجدد خوشه‌بندی می‌کنیم تا دقت خوشه افزایش یابد. برای این کار از الگوریتم خوشه‌بندی K-means استفاده می‌کنیم. اساس کار این الگوریتم بر اساس حداقل فاصله داده‌ها از مرکز خوشه‌هایی که به‌صورت تصادفی و از بین داده‌های موجود انتخاب می‌شود، عمل خوشه‌بندی را انجام می‌دهد. گام‌های این الگوریتم به‌صورت زیر می‌باشد:

- ابتدا k میانگین یعنی، $(\mu_1^{(0)}, \mu_2^{(0)}, \mu_3^{(0)}, \dots, \mu_k^{(0)})$ را نماینده خوشه‌ها هستند، به صورت تصادفی مقدار دهی می‌کنیم.

سپس، دو مرحله زیر را چندین بار اجرا می‌کنیم تا میانگین‌ها به یک ثبات کافی برسند و یا مجموع واریانس های خوشه‌ها تغییر چندانی نکنند [۱۷،۱۸]:

- از میانگین‌ها k خوشه می‌سازیم، خوشه i ام در زمان t تمام داده‌هایی هستند که از لحاظ اقلیدسی کمترین فاصله را با میانگین $\mu_i^{(t)}$ یعنی میانگین i ام در زمان t دارند. خوشه i ام در زمان t برابر خواهد بود با [۱۹] :

$$(\ \mu_i^t\|^2 \leq \| x_p - \mu_j^{(t)}\|^2\ \forall j, 1 \leq j \leq k\ )$$ رابطه (۵)

- حال میانگین‌ها را بر اساس خوشه‌های جدید به این شکل بروز می‌کنیم:

$$\mu_i^{t+1} = \frac{1}{|S_i^t|}\sum_{x_j \epsilon s_i^t} X_j$$ رابطه (۶)

۴-۶- الگوریتم Apriori

در مرحله بعد به‌منظور شناسایی مجموعه آیتم‌های بزرگ از الگوریتم Apriori استفاده می‌کنیم. روش استفاده از این الگوریتم به‌صورت زیر می‌باشد.

ابتدا مجموعه آیتم‌های تک عضوی که در شرط حداقل پوشش صدق می‌کند را انتخاب می‌کنیم.

با استفاده از مجموعه آیتم‌های تک عضوی، مجموعه آیتم‌هایی با ۲ عضو را می‌سازد. سپس شرط حداقل پوشش را برای مجموعه آیتم‌های تولیدشده در این مرحله بررسی می‌کند و هرکدام از مجموعه آیتم‌های تولیدشده در این مرحله در شرط حداقل پوشش صدق نکند، حذف می‌شود [۲۰، ۲۱].

عمل انجام‌شده در مرحله دوم را برای مجموعه آیتم‌های تولیدشده در مرحله دوم که در شرط حداقل پوشش صدق کردند تکرار می‌کند با این تفاوت که در این مرحله باید به این نکته توجه شود که مجموعه‌های سه عضوی ساخته‌شده نباید شامل مجموعه آیتم‌هایی باشد که در مرحله دوم حذف شدند. مرحله سوم به همین صورت تا تولید بزرگ‌ترین آیتم‌های داری شرط حداقل پوشش، تکرار می‌شود. بدیهی است اگر در مرحله ایی هیچ یک از آیتم‌های داری شرط حداقل پوشش صدق نکردند، الگوریتم متوقف می‌شود [۲۲،۲۳].

پس از شناسایی مجموعه آیتم‌های بزرگ، نوبت استخراج قوانین وابستگی داری حداقل اطمینان موردنظر می‌رسد.

۴-۷- قواعد وابستگی

در مرحله بعد باید روابط و وابستگی‌های میان تراکنش‌های مربوط به خرید مشتریان را پیدا کنیم. برای انجام این کار از قوانین وابستگی استفاده می‌کنیم. قوانین وابستگی، مطالعه صفات یا ویژگی‌هایی می‌باشد که با همدیگر وابستگی زیادی دارند. روش‌های تحلیل وابستگی، وابستگی بین مشخصه‌ها را کشف می‌کنند، به این معنی که آن‌ها قواعدی را پیدا می‌کنند که با استفاده از آن‌ها بتوان روابط بین یک یا چند مشخصه را کمی سازی نمود. همچنین قواعد وابستگی روابط و وابستگی‌های متقابل بین مجموعه بزرگی از اقلام داده ایی را نشان می‌دهند[۲۴، ۲۵]. روش انجام کار بدین‌صورت می‌باشد که مشخصات خریداری شده توسط هر مشتری در یک رکورد پایگاه داده ذخیره می‌شود. به هر رکورد یک شناسه (TID) نسبت داده می‌شود. فرض بر این صورت می‌باشد که مجموعه I شامل تمام آیتم‌های فروشگاه‌ها می‌باشد. اگر مجموعه آیتم‌های $x, y \subset I$ باشد به طوری که $x \cap y = \emptyset$ باشد، آنگاه $x \rightarrow y$ یک قانون وابستگی است که بیان می‌کند اگر مشتری اجناس مجموعه x را خریداری کرده باشد، اجناس مجموعه y را نیز خواهد خرید. که از طریق این روش می‌توان استراتژی فروش، بخش‌بندی مشتریان، تنظیم کاتالوگ‌ها را انجام داد[۲۹، ۳۰، ۳۱].

فرض کنید $I = \{i_1, i_2, i_3, \dots, i_m\}$ مجموعه آیتم‌های موجود باشد شامل m عضو است. به هر زیرمجموعه از مجموعه آیتم‌های i نیز یک مجموعه آیتم می‌گوییم.

مجموعه آیتم $X \subseteq I$ را در نظر بگیرید، اگر تعداد مشتریانی که اقلام X را خریداری کرده‌اند برابر ℓ باشد آنگاه آن را به‌صورت رابطه زیر نشان می‌دهیم.

رابطه (۷) $$P(x) = \ell$$

فرض کنید دو مجموعه آیتم Yو X که زیرمجموعه‌های مجموعه آیتم I هستند را داشته باشیم. قانون وابستگی تولیدشده وسط این دو مجموعه آیتم به‌صورت رابطه $X \Longrightarrow Y$ نمایش داده می‌شود.

مجموعه I شامل تمام اقلام و مجموعه آیتم‌های x و y را در نظر بگیرید. پوشش قانون $X \Longrightarrow Y$ به‌صورت رابطه زیر تعریف می‌شود.

رابطه (۸) $$\boldsymbol{Support\ (X \Longrightarrow Y) = S(X \Longrightarrow Y) = \frac{p(x\ \&\ y)}{m}}$$

که در آن منظور از $p(x\ \&\ y)$ تعداد مشتریانی است که مجموعه اقلام X و Y را با هم خریداری کرده اند و m تعداد کل مشتریان می‌باشد. فاکتور پوشش بیانگر نسبت تعداد مشتریانی که مجموعه اقلام X و Y را باهم خریداری کرده اند به کل مشتریان است.

مجموعه I شامل تمامی اقلام و مجموعه آیتم‌های Xو Y را در نظر بگیرید. درجه اطمینان قانون $X \Longrightarrow Y$ به صورت رابطه زیر تعریف می‌شود [۲۶،۲۷].

رابطه (۹) $$\boldsymbol{Confidence\ (X \Longrightarrow Y) = C\ (X \Longrightarrow Y) = \frac{p(X\&\ Y)}{P\ (X)}}$$

فاکتور درجه اطمینان بیانگر نسبت تعداد مشتریانی است که مجموعه اقلام X و Y را باهم خریداری کرده اند به تعداد مشتریانی که اقلام X را خریداری کرده اند.

بدین‌صورت می‌توان برای تمامی قوانین وابستگی تولیدشده فاکتورهای پوشش و اطمینان را محاسبه کرد. هدف کلی در کشف قوانین وابستگی استخراج قانون‌هایی از جداول اطلاعات است که داری میزان پوشش و اعتبار قابل‌قبول باشند. که از طریق این اطلاعات می‌توان هم نیاز مشتریان هم نیاز مدیران فروشگاه‌ها را تشخیص داد.

بهینه سازی اطلاعات با استفاده از الگوریتم فاخته

در مرحله بعد با استفاده از الگوریتم فرا ابتکاری فاخته داده‌های خود را بهینه خواهیم کرد. الگوریتم فاخته یکی از جدیدترین و قوی‌ترین روش‌های بهینه‌سازی تکاملی است. این الگوریتم از نحوه زندگی پرنده ایی به نام فاخه بهره گرفته است. این الگوریتم با یک جمعیت اولیه کار

خود را شروع می‌کند . این جمعیت از فاخته، تعدادی تخم دارند که آن‌ها را در لانه‌ی تعدادی پرنده میزبان خواهند گذاشت. تعدادی از این تخم‌ها شباهت بیشتری به تخم‌های پرنده میزبان دارند، شانس بیشتری به تخم‌های پرنده میزبان دارند، شانس بیشتری برای رشد و تبدیل‌شدن به فاخته بالغ را خواهند داشت. سایر تخم‌هایی که توسط پرنده میزبان شناسایی‌شده از بین می‌روند. میزان تخم‌های رشد کرده مناسب بودن لانه‌های آن منطقه را نشان می‌دهند. هرچه تخم بیشتری در یک ناحیه، قادر به زیست باشند و نجات یابند به همان اندازه، سود بیشتری به آن منطقه اختصاص داده می‌شود؛ بنابراین موقعیتی که در آن بیشترین تعداد تخم نجات یابند، پارامتری خواهد بود که فاخته‌ها قصد بهینه‌سازی آن رادارند [۳۴، ۳۳، ۳۲].

- برای حل یک مسئله بهینه‌سازی لازم است تا مقادیر متغیرهای مسئله به فرم یک آرایه شکل گیرند.
- در الگوریتم فاخته آرایه را Habitat یا محل سکونت می‌گوییم.
- در مسئله بهینه‌سازی N_{var} بعدی یک Habitat یک آرایه $1 \times var$ خواهد بود که موقعیت فعلی زندگی فاخته‌ها را نشان می‌دهد. این آرایه به شکل زیر تعریف می‌شود.

$$\boldsymbol{Habitat} = [\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, .., \ldots, \boldsymbol{x_{Nvar}}] \quad (۷$$

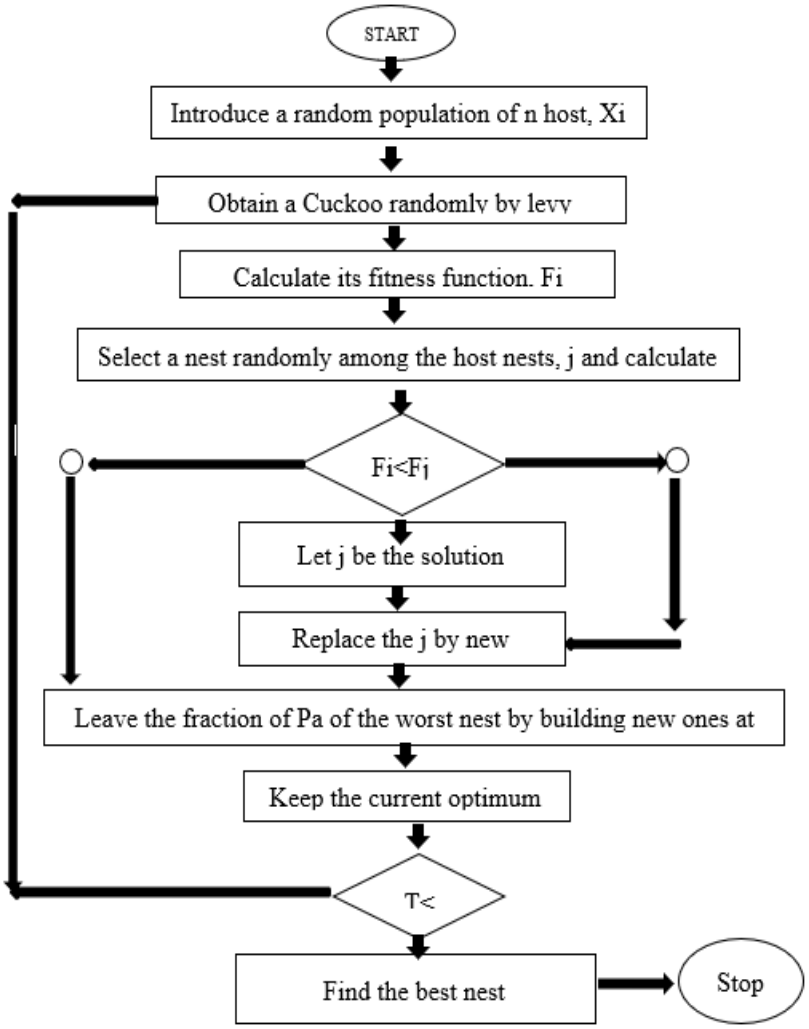
- میزان مناسب بودن در Habitat فعلی با ارزیابی (f_p) در Habitat به دست می‌آید، بنابراین:

$$\boldsymbol{Profit} = \boldsymbol{f_p}\ \boldsymbol{habitat} = \boldsymbol{f_p}(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, .., \ldots, \boldsymbol{x_{Nvar}}) \quad (۸$$

- همان‌طور که مشاهده می‌شود، الگوریتم فاخته الگوریتمی است که تابع سود را ماکزیمم می‌کند.
- برای استفاده از الگوریتم فاخته برای حل مسائل کمینه‌سازی کافی است یک علامت منفی در تابع هزینه ضرب کنیم.
- برای شروع الگوریتم بهینه‌سازی یک ماتریس Habitat با سایز $N_{pop} \times N_{var}$ تولید می‌کنیم.
- سپس برای هرکدام از این Habitat ها تعدادی تخم تصادفی تخصیص می‌یابد.
- در طبیعت هر فاخته بین ۵ تا ۲۰ تخم می‌گذارد. این اعداد به‌عنوان حد بالا و حد پایین تخصیص تخم هر فاخته در تکرارهای مختلف استفاده می‌شوند.
- دیگر عادت فاخته حقیقی این است که آن‌ها را در یک دامنه مشخص تخم‌های خود را می‌گذارند.

$$\boldsymbol{ELR} = \boldsymbol{a} \times \frac{\boldsymbol{Number\ of\ current\ cuckoos\ eggs}}{\boldsymbol{Total\ number\ of\ eggs}} \times (\boldsymbol{Var_{hi}} - \boldsymbol{Var_{low}}) \quad (۹$$

در شکل (۲) چهارچوب الگوریتم فاخته را مشاهده می‌کنید.



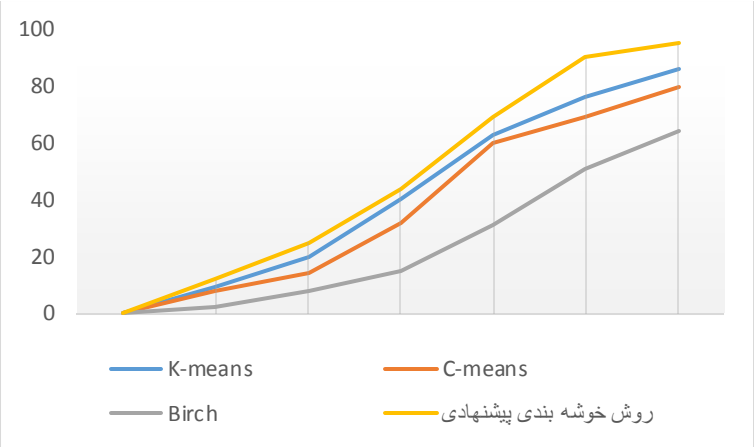
شکل ۲. نمایی از الگوریتم فاخته

۴-۸- توصیه به کاربر

در مرحله آخر می‌توان از اطلاعات به‌دست‌آمده جهت شناسایی و توصیه به کاربران استفاده نمود. می‌توان برای انجام این کار از یک سیستم توصیه گر استفاده کرده و آیتم‌های موردعلاقه کاربر را در اختیار آن‌ها قرارداد. همچنین می‌توان اطلاعات را در اختیار مدیران فروشگاه‌های زنجیره ایی و فروشگاه‌های آنلاین قرارداد تا از این اطلاعات برای بهبود و تشخیص نیازهای کاربران استفاده کنند و محصولات بیشتر خریداری‌شده را کشف و این محصولات را در اختیار کاربران قرار دهند.

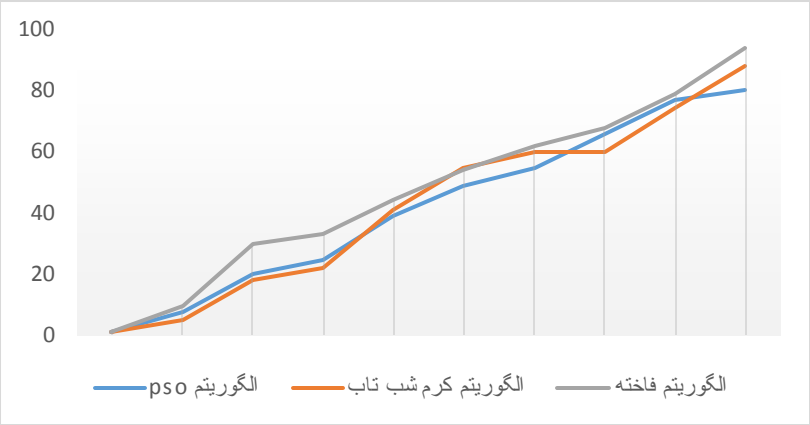
۵. ارزیابی روش پیشنهادی

برای ارزیابی روش پیشنهادی مقایسه‌ای بین الگوریتم خوشه‌بندی استفاده شده در روش پیشنهادی و سایر روش‌های موجود صورت گرفت که نتایج آن به‌صورت زیر می‌باشد. ارزیابی براساس میزان دقت می‌باشد. نتایج بدین صورت می‌باشد که الگوریتم ترکیبی حاصل از ترکیب دو الگوریتم K-Means و C-means توانست میزان دقت ۹۵٪ را به دست آورد این در حالی است که هرکدام از این الگوریتم‌ها به تنهایی درصد دقت کمتری را به دست آوردند به طوری که الگوریتم K-Means به تنهایی درصد دقت ۸۶٪ والگوریتم C-Means توانست درصد دقت ۸۰٪ را بدست آورد و الگوریتم Birch نیز میزان دقت ۶۴٪ ار بدست آورد. با توجه به نتایج ارزیابی بدست آمده می‌توان نتیجه گرفت که الگوریتم ترکیبی روش پیشنهادی توانسته خوب عمل کرده و درصد دقت بهتری را نسبت به سایر روشه ای موجود بدست آورد.



شکل۳. مقایسه میزان دقت الگوریتم خوشه بندی ترکیبی موجود در روش پیشنهادی با سایر روش‌ها

در قسمت ارزیابی الگوریتم‌های بهینه سازی نیز مقایسه ایی بین الگوریتم فاخته و سایر روش‌های موجود صورت گرفت. نتایج حاصل از این مقایسه حاکی در میزان بهینه سازی این بالاتر الگوریتم فاخته نسبت به سایر الگوریتم‌های موجود بود. به طوری که الگوریتم فاخته توانست میزان بهینه سازی ۹۴٪ را به دست آورد این درحالی است که الگوریتم PSO میزان بهینه سازی ۸۰٪ و الگوریتم کرم شب تاب میزان بهینه سازی ۸۸٪ را به دست آورد. با توجه به نتایج به دست آمده می‌توان گفت که الگوریتم فاخته عملکرد خوبی را بدست آورده و نسبت به سایر الگوریتم‌های موجود عملکرد بهتری را بدست آورده است.



شکل۴. مقایسه میزان بهینه سازی الگوریتم فاخته با سایر الگوریتم‌های فراابتکاری

۶- نتیجه گیری

با توجه به نتایج حاصل از روش پیشنهادی ارائه‌شده در داده‌های تحقیق برمی‌آید می‌توان بر اساس سابقه مشتریان و کالاهایی که در سبد خرید خوددارند، اقدام به ارائه توصیه‌های مفید نمود. با توجه به بررسی مدل‌ها مشخص‌شده است که درصورتی‌که سبد خرید مشتری حاوی کالاهای مرتبط با یکدیگر باشند روش پیشنهادی موجود در این تحقیق محصولات مناسب را توصیه می‌کند. از نگاهی دیگر فرایند تجزیه‌وتحلیل و توصیه پیشنهاد خرید محصولی است که در سبد خرید وجود نداشته و در اغلب موارد حتی اگر کاملاً دقیق هم نباشد ممکن است موردتوجه مشتری قرار گیرد و محصولات توصیه‌شده را نیز در کنار محصولات سبد خرید خود قرار دهند، زیرا در برخی موارد مشتریان برخی محصولات موردنیاز خود را فراموش می‌کنند که در این صورت این سیستم به آن‌ها در این زمینه کمک زیادی خواهد نمود، خصوصاً این امر در فروشگاه‌ها که خریدهای زیادی داشته باشند و مشتریان قصد خرید ماهیانه را داشته باشند. در این تحقیق با استفاده از تجزیه‌وتحلیل سبد خرید مشتریان و درک رفتار آن‌ها با توجه به محصولاتی که خریداری کرده‌اند و به هم وابسته‌اند، می‌تواند تصمیمات فروش و بازاریابی متفاوتی را با توجه به سیاست‌های فروش انجام داد. در این تحقیق به ارائه روشی به‌منظور بهبود تحلیل سبد خرید مشتریان فروشگاه‌های زنجیره ایی پرداخته شد که تا حد زیادی می‌تواند تحلیل سبد خرید مشتریان را بهبود بخشد. روش پیشنهادی موجود در این تحقیق استفاده از دو الگوریتم قواعد وابستگی و الگوریتم Apriori است که طبق تحقیقات انجام‌شده روش پیشنهادی موجود در این تحقیق می‌تواند تا حدود ۹۴٪ به‌درستی سبد خرید مشتریان را تحلیل کرده و نیازهای مشتریان و فروشگاه‌ها را به‌درستی تشخیص دهد.

۷-مراجع

[1] Huang, Z. Liu, B. Zhong, Y., Wang, W., An, J., & Jiang, Z. (2021). Data Storage Analysis of Online Shopping Cart Based on WeChat Mini Program. *International Core Journal of Engineering*, *7*(1), 362-365.

[2] Tarhini, A., Alalwan, A. A., Al-Qirim, N., & Algharabat, R. (2018). An analysis of the factors influencing the adoption of online shopping. *International Journal of Technology Diffusion (IJTD)*, *9*(3), 68-87.

[3] Hänninen, M., & Ahlbom, C. P. (2019, July). Dynamic Pricing and Shopping Cart Abandonment in Online Retail: An Abstract. In *Academy of Marketing Science World Marketing Congress* (pp. 263-264). Springer, Cham.

[4] Islam, H., Jebarajakirthy, C., & Shankar, A. (2021). An experimental based investigation into the effects of website interactivity on customer behavior in on-line purchase context. *Journal of Strategic Marketing*, *29*(2), 117-140.

[5] Wu, I. C., & Yu, H. K. (2020). Sequential analysis and clustering to investigate users' online shopping behaviors based on need-states. *Information Processing & Management*, *57*(6), 102323.

[6] Henaff, O. (2020, November). Data-efficient image recognition with contrastive predictive coding. In *International Conference on Machine Learning* (pp. 4182-4192). PMLR.

[7] Nazabal, A., Olmos, P. M., Ghahramani, Z., & Valera, I. (2020). Handling incomplete heterogeneous data using vaes. *Pattern Recognition*, *107*, 107501.

[8] Cuong, D. T. (2020). The effect of brand identification and brand trust on brand commitment and brand loyalty at shopping malls. *International Journal of Advanced Science and Technology*, *29*(7), 695-706.

[9] Giao, H., Vuong, B., & Quan, T. (2020). The influence of website quality on consumer's e-loyalty through the mediating role of e-trust and e-satisfaction: An evidence from online shopping in Vietnam. *Uncertain Supply Chain Management*, *8*(2), 351-370.

[10] Lam, I. K. V., & Wong, I. A. (2020). The role of relationship quality and loyalty program in tourism shopping: a multilevel investigation. *Journal of Travel & Tourism Marketing*, *37*(1), 92-111.

[11] Rashid, M. (2021). A systematic approach of data preparation for ABET accreditation. *Int. J. Eng. Educ*, *37*(1), 1-13.

[12] Yang, L., Zhang, J., Wang, X., Li, Z., Li, Z., & He, Y. (2021). An improved ELM-based and data preprocessing integrated approach for phishing detection considering comprehensive features. *Expert Systems with Applications*, *165*, 113863.

[13] Zhou, Q., & Ooka, R. (2021). Influence of data preprocessing on neural network performance for reproducing CFD simulations of non-isothermal indoor airflow distribution. *Energy and Buildings*, *230*, 110525.

[14] Nida, N., Irtaza, A., Javed, A., Yousaf, M. H., & Mahmood, M. T. (2019). Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. *International journal of medical informatics*, *124*, 37-48.

[15] Hashemzadeh, M., Oskouei, A. G., & Farajzadeh, N. (2019). New fuzzy C-means clustering method based on feature-weight and cluster-weight learning. *Applied Soft Computing*, *78*, 324-345.

[16] Xu, P., Babanezhad, M., Yarmand, H., & Marjani, A. (2020). Flow visualization and analysis of thermal distribution for the nanofluid by the integration of fuzzy c-means clustering ANFIS structure and CFD methods. *Journal of Visualization*, *23*(1), 97-110.

[17] Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, *8*, 80716-80727.

[18] Jahwar, A. F., & Abdulazeez, A. M. (2020). Meta-Heuristic Algorithms For K-Means Clustering: A Review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, *17*(7), 12002-12020.

[19] Yang, W., Long, H., Ma, L., & Sun, H. (2020). Research on clustering method based on weighted distance density and k-means. *Procedia Computer Science*, *166*, 507-511.

[20] Panjaitan, S., Amin, M., Lindawati, S., Watrianthos, R., Sihotang, H. T., & Sinaga, B. (2019, August). Implementation of Apriori Algorithm for Analysis of Consumer Purchase Patterns. In *Journal of Physics: Conference Series* (Vol. 1255, No. 1, p. 012057). IOP Publishing.

[21] Raj, S., Ramesh, D., & Sethi, K. K. (2021). A Spark-based Apriori algorithm with reduced shuffle overhead. *The Journal of Supercomputing*, *77*(1), 133-151.

[22] Zhu, S. (2019). Research on data mining of education technical ability training for physical education students based on Apriori algorithm. *Cluster Computing*, *22*(6), 14811-14818.

[23] Singh, P. K., Othman, E., Ahmed, R., Mahmood, A., Dhahri, H., & Choudhury, P. (2021). Optimized recommendations by user profiling using apriori algorithm. *Applied Soft Computing*, 107272.

[24] Vignoud, G., Venance, L., & Touboul, J. D. (2018). Interplay of multiple pathways and activity-dependent rules in STDP. *PLoS computational biology*, *14*(8), e1006184.

[25] Wong, K. T., Morris, Z. N., & Nnonyelu, C. J. (2019). Rules-of-thumb to design a uniform spherical array for direction finding—Its Cramér–Rao bounds' nonlinear dependence on the number of sensors. *The Journal of the Acoustical Society of America*, *145*(2), 714-723.

[26] Bolin, D., & Wallin, J. (2019). Scale dependence: Why the average CRPS often is inappropriate for ranking probabilistic forecasts. *arXiv preprint arXiv:1912.05642*.

[27] Oh, J., & Weiß, C. H. (2020). On the individuals chart with supplementary runs rules under serial dependence. *Methodology and Computing in Applied Probability*, 1-17.

[۲۸]ئویسی, هادی, غروی, & عرفانه. (۲۰۱۵). رویکرد مبتنی بر داده کاوی در مدیریت ارتباط با مشتری و بازاریابی. *مطالعات رفتار مصرف‌کننده*, ۲ (۲), ۲۵-۴۴.

[29] Yang, T. C., & Lai, H. (2006). Comparison of product bundling strategies on different online shopping behaviors. *Electronic Commerce Research and Applications*, *5*(4), 295-304.

[30] Chiang, W. Y. (2011). To mine association rules of customer values via a data mining procedure with improved model: An empirical case study. *Expert Systems with Applications*, *38*(3), 1716-1722.

[31] Suchacka, G., & Chodak, G. (2017). Using association rules to assess purchase probability in online stores. *Information Systems and e-Business Management*, *15*(3), 751-780.

[32] Boveiri, H. R. (2020). An enhanced cuckoo optimization algorithm for task graph scheduling in cluster-computing systems. *Soft Computing*, *24*(13), 10075-10093.

[33] Cai, X., Niu, Y., Geng, S., Zhang, J., Cui, Z., Li, J., & Chen, J. (2020). An under-sampled software defect prediction method based on hybrid multi-objective cuckoo search. *Concurrency and Computation: Practice and Experience*, *32*(5), e5478.

[34] Inci, M., & Caliskan, A. (2020). Performance enhancement of energy extraction capability for fuel cell implementations with improved Cuckoo search algorithm. *International Journal of Hydrogen Energy*.