

توسعه یک سیستم پیشنهاددهنده محصول طراحی مدلی ترکیبی با بهره‌گیری از روش‌های فیلترینگ مشارکت‌محور، کشف قوانین انجمنی، و بخش‌بندی مشتریان

عباس کرامتی^۱ و روشنگ خالقی^{۲*}

^۱ دانشیار دانشکده مهندسی صنایع پردیس دانشکده‌های فنی دانشگاه تهران

^۲ کارشناس ارشد مهندسی صنایع پردیس دانشکده‌های فنی دانشگاه تهران

(تاریخ دریافت ۹۱/۷/۸ - تاریخ دریافت روایت اصلاح‌شده ۱۳۹۳/۴/۲۹ - تاریخ تصویب ۱۳۹۳/۵/۱۱)

چکیده

توسعه روزافزون اینترنت نحوه تعاملات مشتریان و سازمان‌ها را دستخوش تحولات چشمگیری کرده است. یکی از پیامدهای مهم این پدیده پیدایش و گسترش وبگاه‌های تجارت الکترونیکی و افزایش گرایش کاربران به بهره‌گیری از خدمات خریدوفروش برخط است. تنوع خدمات و اقلام عرضه‌شده در این وبگاه‌ها می‌تواند انتخاب محصولات مناسب را برای مشتریان به فرایندی پیچیده و زمان‌بر مبدل کند. سیستم‌های پیشنهاددهنده، با شناسایی ترجیحات مشتریان، آنان را در مواجهه با انبوه اطلاعات یاری و محصولات و خدماتی منطبق سلیقه‌شان به آن‌ها ارائه می‌کند. هدف این پژوهش ارائه مدلی برای توسعه یک سیستم پیشنهاددهنده محصول به مشتریان یک خرده‌فروشی برخط، با بهره‌گیری از مجموعه‌ای از روش‌های داده‌کاوی، است. با استناد به چارچوب پیشنهادی مدل، نخست مشتریان با تکیه بر رویکرد بخش‌بندی مبتنی بر ارزش طول عمر و با لحاظ کردن نسبی ترجیحات، بر اساس مشخصه‌های مدل RFM، خوشه‌بندی می‌شوند. سپس با بهره‌گیری از ساختار پیشنهاددهی دومرحله‌ای پیشنهادهای گوناگون در دو سطح متمایز از رده‌بندی محصول به هر یک از مشتریان هدف ارائه می‌شود. در مرحله نخست، با بهره‌گیری از روش کشف قوانین انجمنی، تراکشن‌های مشتریان هر خوشه در سطح کلاس محصولات بررسی و با شناسایی الگوها و وابستگی‌های پنهان در داده‌ها قوانین پیشنهاددهی معتبر استخراج و لیستی متشکل از کلاس محصولات پیشنهادی به هر یک از مشتریان هدف ارائه می‌شود. در مرحله دوم با بهره‌گیری از رویکرد فیلترینگ مشارکت‌محور و با استناد به خروجی‌های مرحله پیش ترجیحات خرید مشتریان در سطح اقلام کلاس محصولات پیشنهادی شناسایی و لیست نهایی اقلام محصول به هر یک از مشتریان هدف پیشنهاد می‌شود. نتایج اجرای مدل حاکی از آن است که مدل پیشنهادی با بهره‌گیری از یک رویکرد ترکیبی پیشنهاددهی و با کاهش معضلات ناشی از دو پدیده عدم تراکم و مقیاس‌پذیری، در قیاس با مدل‌های مشابهی که با تکیه بر رویکرد سنتی فیلترینگ مشارکت‌محور توسعه می‌یابند، عملکرد مطلوب‌تری دارد.

واژه‌های کلیدی: ارزش طول عمر، خوشه‌بندی، داده‌کاوی فیلترینگ مشارکت‌محور، قوانین انجمنی.

مقدمه

مشتری مد‌ت‌نظر مدیران قرار می‌گیرد. زیرا حفظ مشتریان سازمان می‌تواند به‌مراتب از به‌دست‌آوردن مشتریان جدید دشوارتر باشد [۱]. از ارکان مهم ابقای مشتریان می‌توان به بازاریابی تک‌به‌تک^۲ اشاره کرد. بازاریابی تک‌به‌تک به کمپین‌های بازاریابی شخصی، که از طریق تحلیل و شناسایی و پیش‌بینی تغییرات در رفتار مشتریان توسعه می‌یابد، اشاره می‌کند [۲].

در سالیان اخیر گسترش روزافزون اینترنت سازمان‌ها و مشتریان را به بهره‌گیری از خدمات تجارت الکترونیکی^۱ متمایل ساخته و سهولت برقراری تعاملات مجازی سبب ترویج و رونق فرایند خریدوفروش برخط^۲ میان مشتریان شده است. در فضای رقابتی کسب‌وکارهای برخط، ابقا یا حفظ مشتریان از محورهای مهمی است که در تدوین برنامه‌های حوزه مدیریت ارتباط با

خرید کاربران هر خوشه در سطح کلاس محصولات گوناگون شناسایی و استخراج شد و در مرحله دوم با استناد به خروجی‌های مرحله قبلی و با بهره‌گیری از رویکرد فیلترینگ مشارکت‌محور^۹ لیست نهایی اقلام پیشنهادی به هر یک از مشتریان هدف ارائه شد.

بررسی پیشینه موضوع

در این بخش پژوهش‌های مرتبط با پیشینه موضوع بررسی می‌شود و پس از معرفی اجمالی سیستم‌های پیشنهاددهنده و نحوه عملکرد آن‌ها، رویکردهای بخش‌بندی مبتنی بر ارزش، رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی، و رویکرد فیلترینگ مشارکت‌محور، به منزله اصلی‌ترین رویکردهایی که در توسعه مدل پیشنهادی از آن‌ها استفاده شده، معرفی می‌شود.

سیستم‌های پیشنهاددهنده

کاهش پیوسته هزینه‌های ذخیره‌سازی و پردازش داده‌ها [۹] موجب شده کسب‌وکارهای برخط بتوانند به طیف گسترده‌ای از اطلاعات، که از طریق تعاملات مجازی یا گونه‌های مختلف داده‌های تراکنش‌های مشتریان قابل استخراج است، دست یابند. از طرفی توسعه روزافزون اینترنت مشتریان را با انبوهی از اطلاعات مرتبط با محصولات و خدمات گوناگون مواجه کرده است. این در حالی است که مواجهه با انبوه اطلاعات موجود می‌تواند کاربران را هنگام تصمیم‌گیری در انتخاب محصولات مورد نیازشان با دشواری‌هایی روبه‌رو کند. در چنین وضعیتی ارائه اطلاعات مختصر و در عین حال سودمند به کاربران می‌تواند به تصمیم‌گیری رضایت‌بخش آن‌ها منجر شود و به تبع آن سازمان‌ها نیز علاوه بر افزایش سودآوری خود قادر خواهند بود روابط بلندمدت خود را با کاربران‌شان ادامه دهند [۱۴].

یکی از فناوری‌های رایج در مواجهه با چالش‌های ناشی از حجم بالای اطلاعات فیلترینگ اطلاعات است، که بر طبقه‌بندی جریان محتویات جدید در قالب گروه‌ها تمرکز می‌کند [۱۵]. روش‌های فیلترینگ

به‌کارگیری سیستم‌های پیشنهاددهنده^۴ یکی از شیوه‌های تحقق و اجرای بازاریابی تک‌به‌تک در محیط وبگاه‌های تجارت الکترونیکی است [۳-۶].

سیستم‌های پیشنهاددهنده کسب‌وکارهای برخط را قادر می‌سازد با بهره‌گیری از تاریخچه اطلاعاتی و داده‌های کاربران و با شناسایی ترجیحات آنان اقلام و خدمات شخصی‌سازی شده^۵ را به مشتریان خود ارائه کنند [۷-۱۰]. منظور از شخصی‌سازی ارائه خدمات فردی به هر یک از مشتریان در جهت شناسایی ترجیحات فردی و رفع نیازمندی‌های آنان است [۱۱]. ارائه سرویس‌های شخصی‌سازی شده به مشتریان وبگاه‌های تجارت الکترونیکی این امکان را می‌دهد تا ویژگی‌های اقلام گوناگون را به‌سهولت بررسی و مقایسه و در قیاس با روش‌های سنتی خرید نیازمندی‌های خود را با صرف مدت زمان کمتری برطرف کنند. علاوه بر این، سیستم‌های پیشنهاددهنده ابزارهایی سودمند در بهبود کیفیت فرایند تصمیم‌گیری در محیط‌های پیچیده‌اند و با فراهم آوردن اطلاعات و پیشنهادهای مفید کاربران را در انتخاب کالاهایی که منطبق با سلیق و نیازمندی‌های آنان است یاری می‌کنند [۱۲]. به‌کارگیری این سیستم‌ها در بستر وبگاه‌های تجارت الکترونیکی بدون شک در افزایش سطح رضایت‌مندی مشتریان و ارتقای میزان وفاداری آنان در قبال سازمان مؤثر است و در نهایت تأثیری مطلوب بر سودآوری این کسب‌وکارها خواهد داشت [۱۳].

هدف این پژوهش به‌کارگیری مجموعه‌ای از روش‌های داده‌کاوی^۶ در زمینه طراحی و توسعه مدلی جدید برای ارائه پیشنهادهای کارآمد و منطبق بر ترجیحات به کاربران یک وبگاه تجارت الکترونیکی است. در این مدل رویکرد بخش‌بندی مبتنی بر ارزش^۷ در کنار ترکیبی از رویکردهای پیشنهاددهی برای ارائه پیشنهاد در دو سطح گوناگون رده‌بندی محصول مورد نظر به کار رفت. نخست مشتریان، بر اساس مشخصه‌های ارزش طول عمر و با مد نظر قرار دادن ترجیحات خرید، خوشه‌بندی شدند و در ادامه پیشنهاددهی در دو مرحله انجام گرفت. در مرحله اول، با بهره‌گیری از روش کشف قوانین انجمنی^۸ الگوهای

قالب طیف گسترده‌ای از زمینه‌های کاربردی، از موسیقی و فیلم گرفته تا محتویات صفحات وب، به صورت‌های گوناگون به کاربران ارائه می‌شود. از شیوه‌های رایج ارائه این خروجی‌ها به کاربران تنظیم لیستی از Top-N قلم پیشنهادی است [۳-۵].

به طور کلی، رویکردهای پیشنهاددهی در قالب سه گروه اصلی قابل دسته‌بندی‌اند [۷]: فیلترینگ مشارکت‌محور، فیلترینگ محتوای‌محور^{۱۷}، رویکردهای پیشنهاددهی تلفیقی^{۱۸}. سیستم‌های مبتنی بر رویکرد فیلترینگ محتوای‌محور از طریق تطبیق پروفایل مشتریان (علاقه‌مندی‌های آنان) با ویژگی‌های محتوایی، مانند ویژگی‌های محصولات، محصول مناسب را به مشتری پیشنهاد می‌کند [۴، ۲۰، ۲۱]. مهم‌ترین محدودیت این رویکرد معضل تجزیه و تحلیل محدود محتوایی است. زیرا این رویکرد محدود به ویژگی‌های اقلامی است که پیشنهاد داده می‌شود. به عبارتی، این روش برای آن دسته از کالاهایی که محتوای آن‌ها به صورت الکترونیکی قابل دسترسی نیست یا محتوای آن‌ها مبتنی بر فاکتورهای ذهنی^{۱۹}، از قبیل کیفیت و مد و پسند افراد است، کارایی ندارد [۲۲]. اما رویکرد فیلترینگ مشارکت‌محور بر اساس نظر و ترجیح کاربرانی که با استناد به الگوی سبد خریدشان بیشترین مشابهت را با کاربر مورد هدف دارند پیشنهاددهی مناسب را به وی ارائه می‌کند [۵، ۲۳-۲۵]. در این روش به منظور تعیین تشابهات ذهنی میان کاربران معیارهای تشابه^{۲۰} بین رتبه‌بندی ترجیحات آنان محاسبه می‌شود [۲۶] و در ادامه سیستم می‌کوشد هم‌تایان^{۲۱} هر یک از کاربران هدف را شناسایی کند. منظور از هم‌تایان یک کاربر هدف سایر کاربرانی است که سلاقی مشابه با وی دارند یا افرادی که رتبه‌دهی‌های اختصاص‌یافته از جانب آن‌ها به اقلام گوناگون بیشترین مشابهت را با رتبه‌دهی‌های کاربر مورد نظر دارد. معضل شروع سرد^{۲۲} از محدودیت‌های مهمی است که سیستم‌های مبتنی بر رویکرد فیلترینگ مشارکت‌محور با آن مواجه‌اند. این معضل، که با نام کاربران جدید یا اقلام جدید نیز شناخته می‌شود، زمانی بروز می‌کند که محصول یا فردی جدید به سایت وارد می‌شود [۷، ۲۴].

اطلاعات ترجیحات کاربران را تجزیه و تحلیل و به کاربران وبگاه‌های تجارت الکترونیکی در فرایند انتخاب محصولات مناسب کمک می‌کنند. در واقع این روش‌ها، با فیلتر کردن اطلاعاتی که مشتریان فراهم کرده‌اند، می‌کوشند رفتار خرید کاربران را دنبال و محصولات مناسبی را به آنان پیشنهاد کنند. سیستم‌های فیلترینگ اطلاعات به پروفایلی از نیازمندی‌ها و ترجیحات کاربران احتیاج دارند و نقش عمده‌ای در توسعه سیستم‌های پیشنهاددهنده ایفا می‌کنند. از اواسط دهه ۱۹۹۰ میلادی، که نخستین پژوهش‌ها در حوزه روش‌های فیلترینگ مشارکت‌محور انجام شد، پژوهش در حوزه سیستم‌های پیشنهاددهنده به یکی از زمینه‌های تحقیقاتی پرتعداد تبدیل شد [۱۰، ۱۲]. از سیستم‌های پیشنهاددهنده معروف و موفق می‌توان به سایت آمازون^{۱۰} برای پیشنهاددهی کتاب و اقلام دیگر، گروه لنز^{۱۱} برای پیشنهاددهی مقالات خبری، رینگو^{۱۲} برای پیشنهاددهی موسیقی، نت فلیکس^{۱۳} برای پیشنهاددهی انواع دی‌وی‌دی، فب^{۱۴} برای پیشنهاددهی صفحات وب، و نیوزیدر^{۱۵} که نمونه‌ای از یک سیستم پیشنهاددهنده متون است اشاره کرد [۱۲، ۱۶، ۱۷].

در پیشینه موضوع طبقه‌بندی‌های متفاوتی از ابعاد گوناگون سیستم‌های پیشنهاددهنده، کاربردها، و روش‌های توسعه آن‌ها ارائه شده است. سه مؤلفه عمده‌ای که در تشریح پیکره سیستم‌های پیشنهاددهنده درباره آن‌ها بحث شده عبارت‌اند از منابع ورودی سیستم، فرایند تولید پیشنهاد، و خروجی‌های سیستم [۸، ۱۳، ۱۸، ۱۹]. منابع ورودی عموماً اطلاعاتی از قبیل ترجیحات مشتریان در قبال اقلام گوناگون، ویژگی‌ها یا رتبه‌دهی‌های اقلام^{۱۶}، و سوابق خرید آنان را شامل می‌شود. سیستم پیشنهاددهنده از این اطلاعات در شناسایی علایق و ترجیحات مشتریان استفاده می‌کند. فرایند پیشنهاددهی، با بهره‌گیری از روش‌ها و رویکردهای مشخص، داده‌های ورودی به سیستم را جهت تولید لیست نهایی پیشنهادها به کار می‌گیرد. این فاز مهم‌ترین نقش را در مسیر توسعه سیستم پیشنهادگر ایفا می‌کند. در نهایت، خروجی‌های سیستم، که در

بخش‌بندی مبتنی بر ارزش و کاربرد آن در توسعه سیستم‌های پیشنهاددهنده

بخش‌بندی بازار، که در آن مشتریانی با نیازها و مشخصه‌های مشابه کنار یک‌دیگر و در قالب دسته‌هایی همگن گروه‌بندی می‌شوند، می‌تواند در اجرای موفقیت‌آمیز برنامه‌های مدیریت ارتباط با مشتریان، ارزیابی وفاداری مشتریان، و توسعه استراتژی‌های بهینه کسب‌وکار در قبال هر یک از گروه‌های مشتریان نقشی کلیدی ایفا کند [۲۷-۲۹]. بخش‌بندی مشتریان از رویکردهایی است که بارها محققان در توسعه سیستم‌های پیشنهاددهنده به کار برده‌اند [۵، ۲۱، ۲۳، ۲۵]. انتخاب مشخصه‌های بخش‌بندی در یک مدل پیشنهادگر اهمیت بالایی دارد. در حوزه پژوهشی سیستم‌های پیشنهادگر طیف گوناگونی از مشخصه‌هایی که از طریق پروفایل کاربری مشتریان قابل استخراج‌اند، اعم از داده‌های تصریحی یا تلویحی، در بخش‌بندی کاربران این سیستم‌ها به کار رفته است [۳، ۵، ۲۳]. رویکرد بخش‌بندی مبتنی بر ارزش روشی کارآمد در تجزیه و تحلیل و کشف ویژگی‌های ارزشی مشتریان و شناسایی مشتریان ارزشمند سازمان است. منظور از ویژگی‌های ارزشی مشخصه‌هایی است که در ارزیابی و تعیین ارزش مشتریان یک سازمان اثرگذار است. برخلاف بسیاری از رویکردهای سنتی بخش‌بندی، که در آن‌ها شاخص‌هایی از قبیل متغیرهای جمعیت‌شناختی و روان‌شناختی یا رفتاری برای گروه‌بندی مشتریان به کار می‌رود، رویکرد بخش‌بندی مبتنی بر ارزش مشتریان را بر اساس ارزش طول عمر^{۲۶} و میزان سودآوری آنان گروه‌بندی می‌کند [۵، ۲۷، ۳۰، ۳۱]. از دیدگاه سازمان، ارزش طول عمر مشتری عبارت است از مجموعه‌ای از معاملات انجام گرفته بین سازمان و مشتریان در طول مدت زمانی که یک فرد در جایگاه مشتری با سازمان مربوطه ارتباط تجاری دارد [۳۲]. مدل *RFM* از مدل‌هایی است که در محاسبه ارزش طول عمر و اجرای بخش‌بندی مبتنی بر ارزش کاربرد فراوان دارد [۵، ۲۸، ۳۳]. متغیرهای این مدل عبارت‌اند از ۱. تازگی^{۲۷}، که به مدت زمان سپری‌شده از زمان آخرین خرید مشتری اشاره می‌کند. هر چه این زمان

زمانی که محصولی جدید معرفی می‌شود تا زمانی که تعداد مشخصی از کاربران به آن رتبه ندهند یا آن را نخرند امکان پیشنهاددهی آن از جانب سیستم وجود ندارد. از طرف دیگر، فرایند پیشنهاددهی برای مشتریان جدید که تعداد اقلام خریداری‌شده یا رتبه‌دهی‌شده از سوی آن‌ها محدود یا ناموجود است ناممکن یا دشوار است. به علاوه، معضل مقیاس‌پذیری^{۲۳} و عدم تراکم^{۲۴} از چالش‌های مهمی است که توسعه سیستم‌های پیشنهاددهنده را با محدودیت‌های فراوان مواجه می‌کند [۸، ۹]. در بسیاری موارد، تعداد کاربران و تنوع اقلام عرضه‌شده در وبگاه‌های برخط بسیار بالاست. در چنین مواردی، بروز هزینه‌های بالای محاسباتی در توسعه سیستم‌های پیشنهاددهنده امری اجتناب‌ناپذیر است. زیرا محاسبات الگوریتم‌های پیشنهاددهی با افزایش تعداد مشتریان و اقلام موجود در پایگاه داده به صورت غیر خطی افزایش می‌یابد [۷]. از آنجا که طیف اقلام ارائه‌شده در وبگاه‌های برخط بسیار گسترده است، همپوشانی ترجیحات میان کاربران در اغلب موارد بسیار ناچیز و در مواردی ناموجود است. در چنین مواردی ماتریس‌های کاربر-قلم^{۲۵}، که غالباً استخراج ترجیحات به کمک آن‌ها انجام می‌گیرد، با معضل عدم تراکم روبه‌رو می‌شود و این محدودیت فرایند تولید پیشنهاد را دشوار می‌کند.

توسعه و طراحی رویکردهای پیشنهاددهی مبتنی بر روش‌های تلفیقی راهکاری رایج جهت رفع نقایص و محدودیت‌های سیستم‌های پیشنهاددهنده و نیز بهبود سطح کیفیت پیشنهادهای ارائه‌شده به کاربران است [۴، ۲۱، ۲۵]. به منظور کاهش معضلات ناشی از مقیاس‌پذیری و عدم تراکم، در طراحی مدل این تحقیق یک رویکرد پیشنهاددهی سلسله‌مراتبی به کار رفت که در آن از ترکیب روش‌های فیلترینگ مشارکت‌محور و پیشنهاددهی مبتنی بر کشف قوانین انجمنی برای تولید لیست نهایی پیشنهادها استفاده می‌شود. چارچوب تفصیلی مدل پیشنهادی تحقیق در بخش‌های آتی می‌آید.

کوتاه‌تر باشد احتمال خرید مجدد از جانب مشتری بیشتر است؛ ۲. فراوانی^{۲۸}، که به تعداد خریده‌ها یا تعداد معاملات انجام‌شده از سوی مشتری در یک دوره زمانی مشخص اشاره می‌کند. مقادیر بیشتر این متغیر نشان‌دهنده وفاداری بیشتر مشتری است؛ ۳. ارزش پولی^{۲۹}، که بیانگر میزان پول خرج‌شده از جانب مشتری در یک دوره زمانی مشخص است. مقادیر بالای این متغیر نشان‌دهنده آن است که سازمان باید تمرکز بیشتری بر مشتری مورد نظر داشته باشد.

در سالیان اخیر مقالات متعددی در زمینه بررسی کاربرد مفهوم ارزش طول عمر در توسعه سیستم‌های پیشنهادگر و تجزیه و تحلیل نقش مدل‌های ارزش طول عمر در نحوه عملکرد این سیستم‌ها نگاشته شده‌اند [۵، ۲۳، ۳۴]. لیو و شیه متدولوژی جدیدی برای پیشنهاد محصول به مشتریان ارائه کردند که در آن از روش‌های داده‌کاوی و روش AHP^{۳۰} استفاده شده است [۳۵]. اوزان متغیرهای مدل با بهره‌گیری از روش AHP استخراج و روش خوشه‌بندی برای بخش‌بندی مشتریانی که بر اساس مقادیر متغیرهای مدل WRFM^{۳۱} رفتاری مشابه داشته‌اند اجرا می‌شود. در نهایت وابستگی‌ها و ارتباطات مجموعه‌اقدام خریداری‌شده از سوی مشتریان هر یک از خوشه‌ها با بهره‌گیری از روش کشف قوانین انجمنی استخراج می‌شود. قوانین استخراج‌شده از هر خوشه برای ارائه پیشنهاد مناسب به مشتری هدف آن خوشه به کار می‌رود. نتایج به‌دست‌آمده برتری این رویکرد پیشنهاددهی را بر رویکرد سنتی CF نشان داد. در پژوهشی دیگر لیو و شیه به منظور توسعه یک سیستم پیشنهاددهنده محصول روشی تلفیقی ارائه کردند که به طور هم‌زمان از مزایای مدل مبتنی بر WRFM و CF مبتنی بر ترجیحات بهره می‌گیرد [۳۴]. روش CF مبتنی بر ترجیحات از داده‌های باینری خرید/عدم خرید برای شناسایی ترجیحات استفاده می‌کند. در ادامه، با بهره‌گیری از یک ضریب همبستگی یکپارچه، مشتریان به طور هم‌زمان بر اساس ارزش متغیرهای مدل WRFM و ترجیحات خریدشان گروه‌بندی می‌شوند و رویکردی مشابه با پژوهش پیشین برای

تولید پیشنهادهای نهایی استفاده می‌شود. در ادامه پژوهش قبلی، شیه و لیو با بهره‌گیری از تقاضای مشتریان سیستمی جدید برای پیشنهاد محصول به مشتریان توسعه دادند [۲۵]. سیستم جدید، برای شناسایی مشتریان مشابه، ترجیحات تعمیم‌یافته^{۳۲} را به کار می‌گیرد که با ترکیب تقاضای مشتریان و ترجیحات خرید آنان تعیین می‌شود. در پژوهشی دیگر، لیو و همکاران با ترکیب دو روش فیلترینگ مشارکت‌محور و پیشنهاددهی مبتنی بر کشف قوانین توالی^{۳۳} سیستم پیشنهاددهنده جدیدی توسعه دادند [۵]. سیستم پیشنهادی از مقادیر متغیرهای مدل RFM برای خوشه‌بندی مشتریان در قالب گروه‌هایی با ارزش طول عمر مشابه استفاده می‌کند. نتایج این پژوهش حاکی از آن بود که بهره‌گیری از رویکرد بخش‌بندی مبتنی بر ارزش کیفیت پیشنهادهای رویکرد تلفیقی تحت بررسی را بهبود می‌دهد. البدوی و شهبازی رویکردی برای پیشنهاد محصول به مشتریان یک فروشگاه آن‌لاین ارائه کردند که از مفهوم رده‌بندی محصول^{۳۴}، تکنیک‌های وب‌کاوی، بخش‌بندی مبتنی بر ارزش طول عمر، و داده‌های سبد خرید برای بهبود کیفیت پیشنهادهای تولیدی استفاده می‌کند [۲۳]. در این پژوهش تکنیک رده‌بندی محصولات برای کاهش حجم فضای محاسباتی و تکنیک وب‌کاوی برای رفع معضل عدم تراکم در ماتریس‌های کاربر-قلم به کار رفت. با بهره‌گیری از تکنیک وب‌کاوی داده‌های فراوانی کلیک، فراوانی جای‌گذاری در سبد خرید، و فراوانی خرید به ازای هر یک از مشتریان استخراج می‌شود. در ادامه، این داده‌ها همراه داده‌های RFM آنان، با محاسبه یک ضریب همبستگی ترکیبی، در بخش‌بندی مشتریان به کار می‌رود. در نهایت، با بهره‌گیری از رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی پیشنهادهای نهایی تولید می‌شود.

جمع‌بندی مطالعات انجام‌شده حاکی از آن است که به مدل‌های ارزش طول عمر در توسعه سیستم‌های پیشنهادگر کمتر توجه شده است. از همین رو در تدوین چارچوب مدل پیشنهادی تحقیق تلاش می‌شود نقش به‌کارگیری رویکرد بخش‌بندی مبتنی بر ارزش در

که مرکز آن کمترین فاصله یا به عبارت دیگر بیشترین مشابهت را با مشاهده مورد نظر دارد. پس از تخصیص همه مشاهدات، مشخصه‌های مرکز هر خوشه با میانگین‌گیری از مشخصه‌های مشاهدات تخصیص‌یافته به آن به‌روز می‌شود؛ ۳. گام دوم تکرار می‌شود. و این فرایند تا زمانی ادامه می‌یابد که تابع معیار^{۴۱} هم‌گرا شود. این امر زمانی رخ می‌دهد که دیگر تغییری در مقادیر مراکز خوشه‌ها به وجود نیاید.

از شاخص‌های رایج اعتبارسنجی خوشه‌ها^{۴۲}، که در تعیین تعداد بهینه خوشه‌های یا همان پارامتر K کاربرد دارند، می‌توان به شاخص دیویس-بولدین^{۴۳} [۲۸]، شاخص سیلوئت [۳۹]، روش وارد^{۴۴} [۲۷]، و شاخص دان^{۴۵} [۳۹] اشاره کرد. به‌علاوه در مقالات متعددی از مدل SOM برای تعیین تعداد بهینه خوشه‌ها استفاده شده است [۳۳]. در این تحقیق تعداد بهینه خوشه‌ها با بهره‌گیری از شاخص دان محاسبه شد.

روش کشف قوانین انجمنی

از مدل‌های مهم داده‌کاوی، که به‌دفعات در توسعه سیستم‌های پیشنهاددهنده به کار رفته، مدل‌های وابستگی^{۴۶} و قوانین انجمنی^{۴۷} است [۵، ۲۱، ۲۳، ۲۵، ۴۰]. مدل‌های اجتماع یا وابستگی در پی کشف و شناسایی ارتباطات پنهان میان اقلام گوناگون سبدهای خرید مشتریان‌اند و به‌طور کلی تجزیه و تحلیل سبد بازار و توسعه برنامه‌های مرتبط با فروش محصولات جانبی از زمینه‌های مهم کاربردی این‌گونه مدل‌هاست [۱]. از مثال‌هایی که در تشریح کاربرد استخراج الگوهای رایج می‌توان به آن اشاره کرد تجزیه و تحلیل سبد خرید کاربران است. این فرایند، که در توسعه سیستم‌های پیشنهاددهنده نیز کارایی بالایی دارد، در پی آن است که با شناسایی ارتباط میان اقلام سبد خرید کاربران به تجزیه و تحلیل عادات و رفتارهای خرید آنان و شناسایی مجموعه اقلام رایج^{۴۸} بپردازد. منظور از مجموعه اقلام رایج اقلامی است که کاربران معمولاً همراه یا در پی یکدیگر می‌خرند. مجموعه اقلام رایج شناسایی شده، در ادامه، برای استخراج قوانین انجمنی استفاده می‌شود. شاخص‌های پشتیبان^{۴۹} و اطمینان^{۵۰}

کنار ترکیبی از رویکردهای پیشنهاددهی مدتظر قرار گیرد و تأثیر کاربرد هم‌زمان این رویکردها در سطح عملکرد مدل و کیفیت پیشنهادهای تولیدشده آزمایش و بررسی شود. با استناد به پیشینه پژوهش، در این تحقیق نیز نخست مشتریان بر اساس مقادیر مشخصه‌های RFM خوشه‌بندی شدند و در ادامه با بررسی داده‌های تراکنشی مشتریان هر خوشه پیشنهادهای مناسب به مشتریان این خوشه‌ها ارائه شد [۵، ۲۳، ۳۵].

روش خوشه‌بندی

خوشه‌بندی از روش‌های رایجی است که محققان به منظور بخش‌بندی مشتریان به‌کرات از آن استفاده کرده‌اند [۲۷، ۲۸، ۳۳، ۳۶]. خوشه‌بندی عبارت است از فرایند بخش‌بندی یک جمعیت ناهمگن^{۳۵} در قالب خوشه‌هایی همگن و مشابه^{۳۶} [۳۷]؛ به نحوی که واریانس بین خوشه‌ای ماکزیمم و در عین حال واریانس درون خوشه‌ای مینیمم باشد.

روش‌های متنوعی برای توسعه الگوریتم‌های خوشه‌بندی وجود دارد که از آن‌ها می‌توان به روش‌های افزایی^{۳۷}، روش مبتنی بر مدل^{۳۸}، روش‌های سلسله‌مراتبی^{۳۹}، و روش مبتنی بر چگالی^{۴۰} اشاره کرد [۳۸]. الگوریتم K -means از الگوریتم‌های معروف و قدیمی است که با تکیه بر رویکرد خوشه‌بندی مبتنی بر روش‌های افزایی توسعه یافته است. سهولت اجرا مهم‌ترین مزیت این الگوریتم و از دلایل اصلی رواج کاربرد آن است [۵، ۲۷، ۳۱، ۳۶]. تعداد خوشه‌ها پارامتری است که به منزله ورودی الگوریتم K -means باید از پیش تعیین شود. در این پژوهش از الگوریتم K -means برای خوشه‌بندی مشتریان استفاده شد. عملکرد این الگوریتم به‌طور خلاصه بدین شرح است [۳۸]: ۱. نخست، الگوریتم به‌طور تصادفی K عدد از مشاهدات را انتخاب می‌کند و بدین ترتیب خوشه‌های اولیه شکل می‌گیرند. مقادیر مشخصه‌های این مشاهدات مرکز خوشه‌های اولیه را تشکیل می‌دهند؛ ۲. عملیات بخش‌بندی میان سایر مشاهدات ادامه می‌یابد، به نحوی که هر یک از مشاهدات به خوشه‌ای اختصاص می‌یابد

تشریح است [۴۲]: فرض کنید I مجموعه‌ای از اقلام کالاها و D مجموعه‌ای از تراکنش‌هاست. قوانین انجمنی استخراج شده از طریق این تراکنش‌ها مطابق رابطه ۱ خواهد بود:

$$X \Rightarrow Y \quad \text{such that } X \subset I, Y \subset I, \quad (1) \\ X \cap Y = \emptyset$$

در رابطه ۱ جزء X مقدمه و جزء Y تالی شناخته می‌شود. شاخص پشتیبان قانونی است برای بیان درصدی از تراکنش‌ها که به طور هم‌زمان X و Y را شامل می‌شوند و شاخص اطمینان قانونی است برای بیان آنکه میان تراکنش‌هایی که X را شامل می‌شوند چه تعداد Y را نیز شامل می‌شوند. به عبارت دیگر:

$$\text{support}(X \Rightarrow Y) = \frac{\text{تعداد تراکنش‌های حاضر در } D \text{ که } X \cup Y \text{ را شامل می‌شوند}}{\text{تعداد کل تراکنش‌های حاضر در } D} \quad (2)$$

$$\text{confidence}(X \Rightarrow Y) = P(Y|X) = \frac{\text{تعداد تراکنش‌های حاضر در } D \text{ که } X \cup Y \text{ را شامل می‌شوند}}{\text{تعداد تراکنش‌های حاضر در } D \text{ که } X \text{ را شامل می‌شوند}} \quad (3)$$

و ارزیابی‌های یک کاربر مشخص در قبال یک قلم کالای مشخص بر اساس علایق و ارزیابی‌های سایر کاربرانی که بیشترین مشابهت را با وی دارند تخمین زده می‌شود. الگوریتم نزدیک‌ترین همسایگی از الگوریتم‌های رایج مشابهت‌محور به کاررفته در توسعه سیستم‌های مبتنی بر روش فیلترینگ مشارکت‌محور است [۵، ۶، ۲۳، ۲۴]. در این پژوهش نیز از این الگوریتم برای تهیه لیست پیشنهادهای نهایی برای هر یک از مشتریان هدف استفاده شد. عملکرد این الگوریتم در بخش سوم بیان می‌شود.

جمع‌بندی پیشینه و معرفی نوآوری‌های

پژوهش

در این بخش، با بررسی مقالات و منابع کتابخانه‌ای، سیستم‌های پیشنهاددهنده و چگونگی عملکرد آن‌ها در شناسایی ترجیحات مشتریان و مهم‌ترین رویکردهای توسعه آن‌ها معرفی می‌شود. به علاوه مفاهیمی از قبیل بخش‌بندی مشتریان، ارزش طول عمر مشتریان، و بخش‌بندی مبتنی بر ارزش به همراه کاربرد این مفاهیم

معیارهای سنجش میزان جذابیت قوانین استخراج شده‌اند و به ترتیب بر میزان رواج و درجه اطمینان قوانین کشف شده دلالت می‌کنند. الگوریتم *Apriori* الگوریتمی پرکاربرد در یافتن مجموعه‌اقلام رایج و الگوهای تکرارشونده موجود در داده‌های تراکنشی است [۵، ۲۵، ۴۱]. این الگوریتم بر اساس رویه‌ای تکرارشونده و با مدنظر قراردادن مقادیر آستانه دو شاخص پشتیبان و اطمینان به شناسایی و کشف مجموعه‌اقلام رایج می‌پردازد و در ادامه الگوریتم قوانین انجمنی معتبر را از طریق مجموعه‌اقلام شناسایی شده استخراج می‌کند. فرایند استخراج قوانین انجمنی از طریق مجموعه‌اقلام رایج کشف شده به این صورت قابل

در این پژوهش، در مرحله نخست فاز پیشنهاددهی از رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی برای تعیین کلاس محصولات پیشنهادی به هر یک از مشتریان هدف استفاده شد.

روش فیلترینگ مشارکت‌محور

همان‌طور که اشاره شد، در رویکرد فیلترینگ مشارکت‌محور لیست اقلام پیشنهادی به مشتریان هدف، با در نظر گرفتن آرای سایر مشتریانی که ترجیحات یا سلیقه آن‌ها نزدیک‌ترین مشابهت را به سلیقه وی داشته‌اند، تهیه می‌شود [۵، ۲۴، ۳۵]. الگوریتم‌های مبتنی بر مشابهت^{۵۱}، که در زمره رویکردهای مبتنی بر حافظه^{۵۲} است، در توسعه سیستم‌های پیشنهادگری که با تکیه بر رویکرد فیلترینگ مشارکت‌محور توسعه می‌یابند کاربرد فراوانی دارند [۵، ۹، ۲۵]. فرض اساسی در الگوریتم‌های مبتنی بر مشابهت آن است که کاربرانی که در گذشته ترجیحات و علایق یا ارزیابی‌های مشابه یکدیگر داشته‌اند به احتمال قوی در ارزیابی‌های آینده نیز مشابه یکدیگر عمل خواهند کرد. از این رو ترجیحات

در مسیر توسعه و طراحی مدل‌های پیشنهاددهی، با استناد به پیشینه تحقیقات انجام‌شده، نیز تکنیک‌های خوشه‌بندی، کشف قوانین انجمنی، و نزدیک‌ترین همسایگی همراه الگوریتم‌های توسعه آن‌ها، به منزله اصلی‌ترین رویکردهایی که برای اجرای مدل پیشنهادی تحقیق به کار می‌رود، بررسی می‌شود.

مطالعات انجام‌شده حاکی از آن است که به مدل‌های ارزش طول عمر در توسعه سیستم‌های پیشنهادگر کمتر توجه شده است. از این رو، در تدوین چارچوب مدل پیشنهادی تحقیق تلاش شد نقش به‌کارگیری رویکرد بخش‌بندی مبتنی بر ارزش در کنار ترکیبی از رویکردهای پیشنهاددهی مدنظر قرار گیرد و تأثیر کاربرد هم‌زمان این رویکردها در سطح عملکرد مدل و کیفیت پیشنهادهای تولیدشده آزمایش شود و عملکرد مدل پیشنهادی با عملکرد مدل‌های پیشینه پژوهش، که از رویکرد پایه‌ای مشابه برای توسعه سیستم پیشنهاددهنده خود بهره گرفته‌اند، مقایسه شود [۵، ۲۵، ۳۴، ۳۵].

بخش‌بندی بازار بر اساس مشخصه‌های مدل *RFM*، کشف و استخراج قوانین انجمنی در جهت شناسایی الگوهای پنهان در داده‌های حاصل از تراکنش‌های خرید مشتریان، و پیشنهاددهی اقلام نهایی بر اساس تکنیک فیلترینگ مشارکت‌محور با بهره‌گیری از الگوریتم نزدیک‌ترین همسایگی اصلی‌ترین رویکردهایی است که در طراحی متدولوژی مدل پیشنهادی این تحقیق مدنظر قرار گرفت. تعریف رویکردی جدید برای بخش‌بندی بازار، که می‌کوشد مشخصه‌های ارزشی و ترجیحات خرید کاربران را هنگام خوشه‌بندی آنان مدنظر قرار دهد، در کنار توسعه رویکردی دومرحله‌ای برای تهیه لیست نهایی اقلام پیشنهادی، مهم‌ترین ویژگی‌های مدل طراحی شده است. به بیان دیگر، این سیستم با بهره‌گیری از رویکرد بخش‌بندی مشتریان، که در آن از مدل پایه‌ای *RFM* برای خوشه‌بندی مشتریان با ترجیحات تقریباً مشابه استفاده شده، و با اجرای سلسله‌مراتبی رویکردهای کشف قوانین انجمنی و فیلترینگ مشارکت‌محور امکان ارائه پیشنهاد به مشتریان را در دو سطح گوناگون از رده‌بندی محصول

فراهم کرده است. در جریان اجرای هر یک از مراحل فاز پیشنهاددهی، با بهره‌گیری از شاخص‌های ارزیابی دقت و یادآوری و معیار *F1*، عملکرد مدل پیشنهادی مسئله با عملکرد مدل‌های دیگر، که بر اساس رویکردهای سنتی یا مشابه موجود در پیشینه پژوهش توسعه یافته‌اند، مقایسه شد [۵، ۲۵، ۳۴، ۳۵]. انتظار می‌رود اجرای متدولوژی پیشنهادی در سطح داده‌های تراکنشی مسئله معضلات ناشی از دو پدیده عدم تراکم و مقیاس‌پذیری را تا حدی کاهش دهد و نهایتاً بتواند در بهبود عملکرد مدل پیشنهادی مؤثر باشد.

مهم‌ترین ویژگی‌ها و نوآوری‌های مدل پیشنهادی تحقیق در ادامه خلاصه می‌شود:

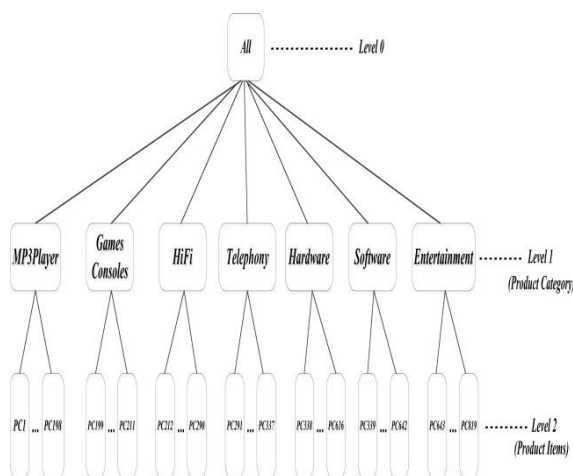
❖ به‌کارگیری هم‌زمان از تکنیک‌های داده‌کاوی: مدل پیشنهادی تحقیق با تکیه بر سه رویکرد خوشه‌بندی و کشف قوانین انجمنی و فیلترینگ مشارکت‌محور توسعه یافت. الگوریتم‌هایی که در اجرای این رویکردها از آن‌ها استفاده شد همگی در زمره الگوریتم‌های رایج و کارآمد حوزه داده‌کاوی‌اند. کاربرد این رویکردها کنار یکدیگر سهولت و دقت فرایند شناسایی و استخراج ترجیحات و سلايق کاربران- فرایندی که در مسیر توسعه سیستم‌های پیشنهاددهنده نقشی کلید دارند- را ارتقا می‌دهد و به بهبود عملکرد سیستم می‌انجامد.

❖ بهره‌گیری از ساختاری تسلسلی در طراحی و توسعه سیستم پیشنهاددهنده: این ساختار به ترتیب امکان تجزیه و تحلیل داده‌های تراکنشی، شناخت ترجیحات کاربران، و ارائه پیشنهادها را در دو سطح گوناگون از رده‌بندی محصول، یعنی کلاس محصولات و اقلام محصول، میسر می‌کند. ساختار پیشنهادی با تفکیک و کاهش فضای محاسباتی اقلام معضلات ناشی از مقیاس‌پذیری^{۵۳} و عدم تراکم داده‌ها را، که از محدودیت‌های شایع در توسعه سیستم‌های پیشنهاددهنده است، تقلیل می‌دهد و زمینه را برای پیشنهاددهی طیف گوناگونی از اقلام متنوع به مشتریان هدف آماده می‌کند.

هدف متدولوژی ارائه‌شده طراحی و توسعه مدلی است که با بهره‌گیری از مجموعه‌ای از روش‌های داده‌کاوی و با تمرکز بر رویکردهای خوشه‌بندی، کشف قوانین انجمنی، و فیلترینگ مشارکت‌محور امکان ارائه پیشنهادهای شخصی را به کاربران یک وبگاه تجارت الکترونیکی فراهم سازد. چارچوب مدل پیشنهادی از چهار فاز اصلی تشکیل می‌شود. در ادامه، مراحل اجرای هر یک از فازهای مدل پیشنهادی بیان می‌شود.

فاز ۱. جمع‌آوری، شناخت، و آماده‌سازی داده‌ها

مدل پیشنهادی تحقیق بر اساس داده‌های یک فروشگاه برخط عرضه‌کننده محصولات دیجیتال - کامپیوتری اجرا شد. اقلام عرضه‌شده در یک فروشگاه می‌توانند در قالب ساختاری سلسله‌مراتبی با سطوح چندگانه گروه‌بندی شوند. این ساختار سلسله‌مراتبی محصولات در بسیاری از فروشگاه‌های خرده‌فروشی موجود است و با نام رده‌بندی محصول^{۵۵} شناخته می‌شود [۴۰ و ۴۳]. بر همین اساس، رده‌بندی محصولات این فروشگاه در دو سطح سلسله‌مراتبی امکان‌پذیر است: کلاس محصولات، که نخستین سطح رده‌بندی محصولات را تشکیل می‌دهد، و اقلام محصول، که مبین سطح دوم رده‌بندی محصولات است. ساختار ساده دوسطحی رده‌بندی محصولات این فروشگاه در شکل ۲ می‌آید.



شکل ۲. ساختار دوسطحی رده‌بندی محصول

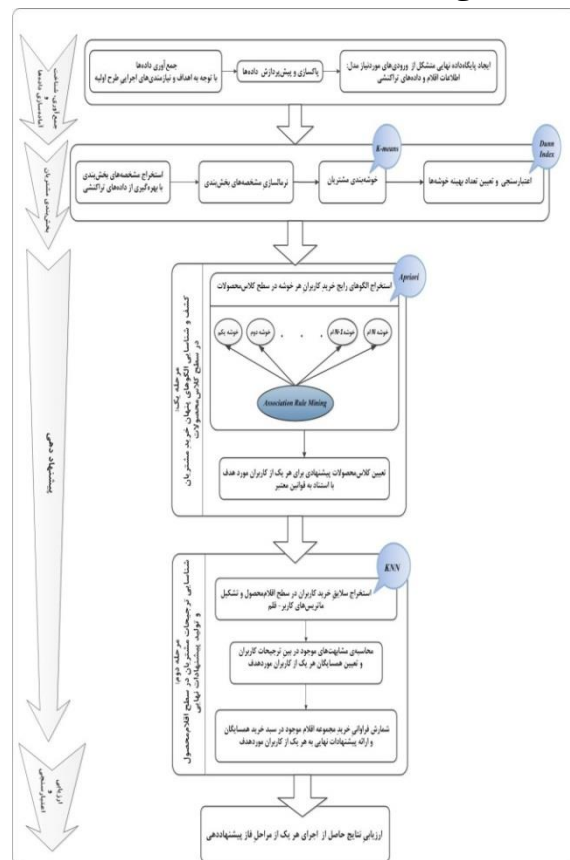
در مرحله پاک‌سازی داده‌ها، داده‌های بی‌کیفیت

طراحی و به‌کارگیری رویکرد توسعه‌یافته‌ای از بخش‌بندی مبتنی بر ارزش: رویکرد بخش‌بندی پیشنهادی تحقیق، که به اختصار $PCB-RFM^{۵۴}$ نامیده شد و در بخش بعدی معرفی می‌شود، با تکیه بر تعاریف پایه‌ای مدل RFM و با لحاظ کردن نسبی ترجیحات و سلاقی مشتریان در جریان اجرای گام خوشه‌بندی، فرایند کشف الگوهای پنهان در تراکنش‌های خرید مشتریان هر خوشه را تسهیل و عملکرد مدل پیشنهاددهنده را بهبود می‌بخشد.

در بخش بعدی چارچوب نهایی مدل پیشنهادی تحقیق و هر یک از فازهای اجرای این مدل به تفصیل معرفی و در بخش چهارم مدل پیشنهادی تحقیق بر داده‌های یک فروشگاه برخط اجرا می‌شود.

معرفی روش‌شناسی و مدل پیشنهادی

شکل ۱ چارچوب تفصیلی مدل پیشنهادی این تحقیق را نمایش می‌دهد.



شکل ۱. چارچوب مدل طراحی‌شده برای سیستم پیشنهاددهنده

حذف شدند. داده‌های بی‌کیفیت داده‌های دو گروه از تراکنش‌ها را شامل می‌شوند؛ نخست تراکنش مشتریانی که تعداد خریدهای آن‌ها از حد آستانه مشخص کمتر است و دوم تراکنش اقلامی که تعداد دفعات خرید آن‌ها از حد آستانه مشخص پایین‌تر است. دلیل حذف این داده‌ها آن است که داده‌های این قبیل مشتریان و اقلام می‌تواند به بروز پدیده عدم تراکم در ماتریس‌های کاربر-قلم منجر شود [۶]. به‌علاوه، کالاهایی که فراوانی خرید پایینی دارند شانس زیادی برای پیشنهاددهی ندارند. فرایند آماده‌سازی داده‌ها، علاوه بر پاک‌سازی، عملیات پیش‌پردازش داده‌ها را نیز دربرمی‌گیرد. برخی از این اعمال عبارت‌اند از نرمال‌سازی داده، استخراج مشخصه‌های مورد نیاز از طریق داده‌های تراکنشی، و تغییر فرم داده‌ها به صورتی که برای اجرای الگوریتم‌های گوناگون آماده شوند.

فاز ۲. بخش‌بندی مشتریان

در این فاز، با بهره‌گیری از الگوریتم k -means، مشتریان بر اساس مقادیر مشخصه‌های RFM و با لحاظ کردن ترجیحات خریدشان در سطح کلاس محصولات گوناگون بخش‌بندی می‌شوند. در این پژوهش برای بخش‌بندی مشتریان رویکردی به نام $PCB-RFM$ تعریف شد. در این رویکرد، با بررسی داده‌های تراکنشی مشتریان در سطح کلاس محصولات، راهکاری مناسب برای اجرای بخش‌بندی مبتنی بر ارزش و استخراج قوانین انجمنی ارائه می‌شود. در واقع، هدف از توسعه رویکرد $PCB-RFM$ آن است که متغیرهای مدل پایه‌ای RFM به نحوی در سطح داده‌های تراکنشی مسئله محاسبه شود که در کنار ارزیابی ارزش طول عمر سلاقی خرید کاربران نیز، از طریق مشخصه‌های محاسبه‌شده، قابل شناسایی و استخراج باشد. اجرای رویکرد $PCB-RFM$ نیازمند آن است که مشخصه‌های RFM مشتریان در سطح هر یک از هفت کلاس محصول موجود به طور جداگانه استخراج شود و در ادامه مجموعه فیله‌های اطلاعاتی محاسبه‌شده به طور هم‌زمان برای خوشه‌بندی مشتریان به کار رود.

فیله‌های خوشه‌بندی در رویکرد $PCB-RFM$ بدین

شرح محاسبه و استخراج می‌شوند: به استثنای متغیر R ، مقادیر متغیرهای FM در سطح هر یک از کلاس محصولات موجود به صورت جداگانه محاسبه می‌شود. به عبارت دیگر، به جای ارزیابی مقادیر متغیرهای FM یک مشتری مفروض در سطح دوم رده‌بندی محصول این متغیرها با تفکیک داده‌های تراکنشی مشتریان در قالب هر یک از کلاس محصولات موجود در سطح یک به طور مجزا ارزیابی می‌شود و مجموعه مشخصه‌های به‌دست‌آمده به مثابه مشخصه‌های بخش‌بندی مشتریان به کار می‌رود. تمایز مشخصه‌های خوشه‌بندی (فیله‌هایی که خوشه‌بندی بر اساس آن‌ها انجام می‌گیرد) برای یک مشتری مفروض در این رویکرد و رویکرد سنتی پیشین در جدول ۱ می‌آید. رویکرد سنتی بخش‌بندی را، که در آن ارزش مشتریان با ارزیابی مقادیر RFM در سطح اقلام همه کلاس محصولات محاسبه می‌شود، RFM کلی یا به اختصار $T-RFM$ می‌نامیم. در جدول ۱، CF_I و CM_I به ترتیب نماد فراوانی خرید و میزان پول خرج‌شده از جانب مشتری مفروض در کلاس محصول I ام است. در واقع، مشخصه‌های CF_I و CM_I به ترتیب صورت تفکیک‌یافته متغیرهای F و M به شمار می‌روند. بدیهی است برای هر مشتری مفروض u روابط ۴ و ۵ برقرار است.

$$\sum_{l=1}^7 CF_{ul} = TF_i \quad (4)$$

$$\sum_{l=1}^7 CM_{ul} = TM_i \quad (5)$$

منطق به‌کارگیری این روش آن است که از میان متغیرهای مدل RFM دو متغیر F و M علاوه بر اینکه در تعیین ارزش مشتریان نقش دارند، می‌توانند منعکس‌کننده ترجیحات خرید آنان نیز باشند. مثلاً، اگر تواتر خرید دو مشتری مفروض در سطح اقلام یک خانواده محصول مشخص به هم نزدیک باشد، می‌توان این‌گونه استدلال کرد که این دو تقریباً به یک میزان به

جدول ۱. معرفی فیلدهای خوشه‌بندی در دو رویکرد *T-RFM* و *PCB-RFM*

RFM over all Product-Items (Total RFM)															
CID	R (days)	F							M (\$)						
01258	12	28							5480						
Product-Category-Based RFM (PCB-RFM)															
CID	R (days)	F							M (\$)						
01258	12	CF ₂	CF ₃	CF ₄	CF ₅	CF ₆	CF ₇	CM ₂	CM ₃	CM ₄	CM ₅	CM ₆	CM ₇		
		5	-	3	12	-	8	-	705	-	945	2450	-	1380	-

منفی است (مقادیر کمتر این مشخصه مطلوب‌تر است)،

با رابطه ۷ قابل اجراست [۵ و ۳۵]:

$$x' = (x - x^S) / (x^L - x^S) \quad (۶)$$

$$(۷)$$

$$x' = (x^L - x) / (x^L - x^S)$$

x و x' به ترتیب نماد مقادیر اصلی و مقادیر نرمالیزه‌شده متغیرهای RFM و x^L و x^S به ترتیب مبین بزرگ‌ترین و کوچک‌ترین مقادیر RFM میان همه مشتریان است. بخش‌بندی مشتریان بر اساس مقادیر نرمال‌سازی‌شده متغیرهای RFM انجام می‌شود.

در آخرین گام فاز بخش‌بندی به ازای مقادیر گوناگون پارامتر K ، که مبین تعداد خوشه‌هاست، فرایند خوشه‌بندی تکرار می‌شود. در هر تکرار با بهره‌گیری از شاخص دان خوشه‌های به‌دست‌آمده اعتبارسنجی و در نهایت تعداد بهینه خوشه‌ها با استناد به تکراری تعیین می‌شود که در آن مطلوب‌ترین مقدار شاخص دان به دست آمده است.

فاز ۳. پیشنهاددهی

در این تحقیق، به منظور استخراج هر چه بهتر روندهای خرید مشتریان در سطح کلاس محصولات و نیز شناسایی ترجیحات آن‌ها در سطح اقلام نهایی، رویکردی دومرحله‌ای، با بهره‌گیری از ترکیب دو روش کشف قوانین انجمنی و فیلترینگ مشارکت‌محور، برای تولید پیشنهاد نهایی طراحی شد. دو مرحله تشکیل‌دهنده فاز پیشنهاددهی بدین شرح است:

مرحله نخست: شناخت الگوهای رایج خرید در

سطح کلاس محصولات

در گام اول، روش کشف قوانین انجمنی، با بهره‌گیری از

خرید از اقلام این گروه محصول تمایل داشته‌اند. فیلد متغیر M نیز به همین صورت قابل استدلال است. چنانچه مقدار پول خرج‌شده از جانب دو مشتری مفروض در سطح یک کلاس محصول مشخص نزدیک هم باشد، می‌تواند بیان‌کننده آن باشد که این دو مشتری تقریباً به میزانی مشابه به تأمین مایحتاج خود از طریق اقلام این کلاس محصول گرایش دارند. به‌علاوه، از رویکردهایی که در پیشینه پژوهش به دفعات برای نمایش ترجیحات مشتریان به کار رفت بهره‌گیری از ماتریس کاربر-قلم بود که در آن داده‌های باینری خرید/عدم خرید مشتریان در قالب مجموعه‌ای از صفر و یک‌ها نمایش داده می‌شود. در موارد متعدد داده‌های این ماتریس مبنای خوشه‌بندی مشتریان در قالب گروه‌هایی بود که از نظر ترجیح خرید به یکدیگر شبیه بودند [۲۵، ۳۴]. متدولوژی پیشنهادی مدل فعلی برای بخش‌بندی مشتریان را می‌توان حالت توسعه‌یافته‌ای از این رویکرد دانست. در واقع، فیلدهای متغیر F مبین سطریهای ماتریس مشتری-کلاس محصولی است که در آن به جای به‌کارگیری داده‌های باینری خرید/عدم خرید از داده‌های غیر باینری فراوانی خرید استفاده شده است.

مشخصه‌های بخش‌بندی، پس از استخراج، باید نرمال‌سازی شوند. پس از نرمال‌سازی داده‌ها هر یک از مشخصه‌های محاسبه‌شده به عددی در بازه صفر تا یک تبدیل می‌شوند. نرمال‌سازی مشخصه‌های RFM برای دو مشخصه F و M که در تعیین ارزش طول عمر و وفاداری مشتریان تأثیر مثبت دارند (مقادیر بزرگ‌تر این دو مشخصه مطلوب‌تر است)، با رابطه ۶ و برای مشخصه R ، که اثرگذاری آن در تعیین ارزش طول عمر مشتریان

استخراج می‌شوند. خروجی قوانین اکتشافی، که همان گروه محصولات پیش‌بینی شده برای هر یک از مشتریان هدف است، ورودی‌های مرحله دوم فرایند پیشنهاددهی خواهند بود.

مرحله دوم: استخراج ترجیحات در سطح اقلام محصول و تولید پیشنهادهای نهایی

در مرحله دوم فاز پیشنهاددهی، لیست نهایی اقلام پیشنهادی برای هر یک از مشتریان هدف با استناد به نتایج مرحله نخست و با بهره‌گیری از اطلاعات تراکنش‌های خرید آنان در سطح اقلام محصول، بیانگر ترجیحات آنان در سطح دوم رده‌بندی محصول است، تهیه می‌شود. در این مرحله، اجرای رویکرد پیشنهاددهی مبتنی بر فیلترینگ مشارکت‌محور امکان ارائه پیشنهادهای شخصی به هر یک از مشتریان هدف را فراهم می‌کند. این مرحله با بهره‌گیری از الگوریتم نزدیک‌ترین همسایگی اجرا می‌شود. گام‌های اجرایی این الگوریتم به طور خلاصه بدین شرح است [۵، ۲۴، ۲۵، ۳۰، ۳۵]:

۱. استخراج ترجیحات کاربران در سطح اقلام محصول:

با بهره‌گیری از داده‌های تراکنشی، ماتریس‌های کاربر-قلم، که داده‌های باینری خرید/عدم خرید موجود در آن‌ها به طور تلویحی مبین ترجیحات خرید مشتریان در سطح اقلام محصول است، استخراج می‌شوند. فرض کنید ماتریس‌های کاربر-قلم^{۵۶} را با R نمایش می‌دهیم. عنصر r_{ua} از این ماتریس در صورتی مقدار ۱ را به خود اختصاص دهد که مشتری مفروض u کالای a را خریده باشد؛ در غیر این صورت مقدار اختصاص یافته به این درایه برابر ۰ خواهد بود.

۲. تشکیل همسایگی‌ها:

همبستگی پیرسونو با استفاده از داده‌های ماتریس‌های کاربر-قلم، میزان مشابهت ترجیحات خرید هر یک از مشتریان هدف با سایر مشتریان محاسبه می‌شود. در ادامه k عدد از مشتریانی که ترجیحات خریدشان بیشترین مشابهت را با مشتری هدف دارد نزدیک‌ترین همسایگان این مشتری انتخاب خواهند شد. تشابه ترجیحات میان دو مشتری مفروض u و v با محاسبه

داده‌های تراکنشی ماتریس‌های کاربر-کلاس محصول، در سطح هر یک از خوشه‌ها اجرا می‌شود. این رویکرد را «استخراج قوانین انجمنی در سطح کلاس محصولات» می‌نامیم. ساروار و همکاران در پژوهش خود به بررسی و معرفی رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی پرداختند [۲۴]. در گام دوم و پس از استخراج قوانین انجمنی معتبر از رویکردی مشابه برای تعیین لیستی متشکل از $Top-N$ گروه محصول پیشنهادی برای هر یک از مشتریان هدف استفاده می‌شود. برای هر مشتری مفروض u ، مجموعه‌ای متشکل از همه گروه محصولات که وی قبلاً آن‌ها را خریده است، ایجاد می‌شود. نماد این مجموعه X_u است. سپس، همه قوانین استخراج شده به فرم $X \Rightarrow Y$ ، که در آن‌ها رابطه $X \subseteq X_u$ برقرار است، شناسایی می‌شود. این رابطه به معنای آن است که مشتری مفروض u ام از همه کلاس محصولات X حداقل یک قلم کالا خریده است. سپس، با توجه به هر یک از قوانین به دست آمده برای مشتری u ام همه کلاس محصولات مجموعه Y کلاس محصولات کاندید برای پیشنهاددهی به وی در نظر گرفته می‌شوند. اولویت هر یک از کلاس محصولات پیشنهادی به مقدار شاخص اطمینان قانون آن بستگی دارد. اگر کلاس محصولات مورد نظر را چندین قانون پیشنهاد داده باشد، قانونی که بیشترین مقدار شاخص اطمینان را دارد انتخاب می‌شود. اگر مجموعه قوانین استخراج شده برای مشتری مفروض u ام با P_u نشان داده شود، در قدم بعدی کلاس محصولات P_u بر اساس مقادیر شاخص اطمینان مربوطه مرتب می‌شود؛ به نحوی که کلاس محصولاتی که مقدار اطمینان آنان بزرگ‌تر است اولویت بالاتری در پیشنهاددهی دارند. در نهایت $Top - N$ کلاس محصول ابتدای لیست، کلاس محصولی خواهد بود که باید به مشتری u ام پیشنهاد شود.

بنا بر آنچه آمد، در مرحله نخست فاز پیشنهاددهی با اجرای روش کشف قوانین انجمنی الگوهای رایج خرید مشتریان هر خوشه در سطح کلاس محصولات گوناگون با استناد به قوانینی معتبر، که مینیمم دو شاخص اطمینان و پشتیبان را برآورده می‌سازند،

ضریب همبستگی پیرسون با رابطه ۸ اندازه‌گیری می‌شود.

$$Sim_{uv}^{PC} = \frac{\sum_{\alpha \in I} (r_{u\alpha} - \bar{r}_u)(r_{v\alpha} - \bar{r}_v)}{\sqrt{\sum_{\alpha \in I} (r_{u\alpha} - \bar{r}_u)^2 \sum_{\alpha \in I} (r_{v\alpha} - \bar{r}_v)^2}} \quad (۸)$$

نمادهای $\bar{r}_v$ و $\bar{r}_u$ به ترتیب بیانگر میانگین تعداد اقلامی است که مشتری u و v ام خریده‌اند. به علاوه، متغیر I معرف مجموعه اقلام موجود است و r_{va} و r_{ua} نشان‌دهنده وضعیت خرید قلم‌کالای α (خرید/عدم خرید) از جانب مشتری u و v ام.

۳. تهیه لیست پیشنهادی: پس از تشکیل همسایگی‌ها برای هر یک از مشتریان هدف، الگوریتم با تجزیه و تحلیل و شناسایی اقلامی که همسایگان به خرید آن‌ها بیشترین تمایل را داشته‌اند و با بهره‌گیری از رویکرد پیشنهاددهی مبتنی بر شمارش فراوانی خرید اقلام لیست نهایی پیشنهاد به هر یک از مشتریان هدف را ارائه می‌کند. برای یک مشتری مفروض تواتر خرید اقلام گوناگونی که همسایگان وی خریده‌اند شمرده می‌شود. در ادامه، این اقلام بر اساس میزان تواتر خریدشان مرتب می‌شوند و لیست نهایی پیشنهادها، که متشکل از N کالایی است که به ترتیب بیشترین تواتر خرید را داشته‌اند و پیش از این مشتری هدف مفروض آن را نخریده است، به وی ارائه می‌شود.

هر یک از گام‌های فوق باید در سطح هر یک از خوشه‌ها جداگانه اجرا شود. گفتنی است در گام تشکیل همسایگی‌ها داده‌هایی که همسایگی هر یک از مشتریان هدف با استناد به آن‌ها شکل می‌گیرد به داده‌های آن گروه از تراکنش‌هایی محدود می‌شود که با استناد به خروجی‌های مرحله نخست فاز پیشنهاددهی در لیست گروه محصولات پیش‌بینی شده برای مشتری مفروض به ثبت رسیده‌اند.

فاز ۴. ارزیابی و اعتبارسنجی

هدف این فاز ارزیابی خروجی‌های هر یک از مراحل فاز پیشنهاددهی و به عبارت دیگر سنجش عملکرد و بررسی کیفیت پیشنهادهای ارائه شده است. با توجه به رواج به کارگیری دو معیار دقت d_u^{57} و یادآوری d_u^{58} و معیار

F_1 در ارزیابی سیستم‌های پیشنهاددهنده، به‌ویژه پژوهش‌هایی که حوزه کاربرد سیستم پیشنهاددهنده در آن‌ها تقریباً مشابه تحقیق پیش‌روست [۹، ۲۳، ۲۵، ۳۵، ۴۰]، این معیارها برای ارزیابی عملکرد مدل پیشنهادی مسئله فعلی انتخاب شدند.

یک فروشگاه خرده‌فروشی برخط را، مانند آنچه در مسئله فعلی بررسی می‌شود، در نظر بگیرید. اقلام گوناگون عرضه شده در قالب دو گروه اصلی قابل دسته‌بندی‌اند [۵]: اقلامی که کاربر به خرید آن‌ها علاقه‌مند است و اقلامی که کاربر به خرید آن‌ها تمایل ندارد. اقلام پیشنهادی سیستم ممکن است برای کاربر جذاب باشد یا جذاب نباشد. معیار یادآوری به کارایی سیستم در ارائه پیشنهادهای جذاب اشاره می‌کند و معیار دقت نشان‌دهنده آن است که اقلام پیشنهادی در واقعیت به چه میزان از دید کاربر جذاب است. دو معیار یادآوری و دقت به زبان ساده به صورت زیر قابل تعریف است:

$$\text{تعداد اقلامی که به درستی پیشنهاددهی شده‌اند} = \frac{\text{تعداد اقلام جذاب}}{\text{یادآوری}}$$

$$\text{تعداد اقلامی که به درستی پیشنهاددهی شده‌اند} = \frac{\text{تعداد اقلام پیشنهادی}}{\text{دقت}}$$

اقلام جذاب از دید مشتری هدف u ام اقلامی‌اند که وی در طی دوره آزمون می‌خرد. اقلامی که درست پیشنهاد داده شده‌اند اقلامی هستند که با اقلام جذاب خریداری شده از جانب مشتری هدف در دوره آزمون مطابقت دارند و در نهایت اقلام پیشنهادی به تعداد اقلام موجود در لیست پیشنهادها اشاره می‌کند. اگر تعداد اقلامی را که به مشتری هدف u ام درست پیشنهاد شده‌اند با $d_u(N)$ ، تعداد اقلام جذاب از دید مشتری هدف u ام را با D_u ، و تعداد اقلام پیشنهادی را با N نشان دهیم، دو معیار یادآوری و دقت برای هر یک از مشتریان هدف با رابطه ۹ و ۱۰ بازنویسی و محاسبه می‌شود.

$$R_u = \frac{d_u(N)}{D_u} \quad (۹)$$

محیط اجرا و توصیف داده‌ها

نرم‌افزار عمده‌ای بوده‌اند که فازهای گوناگون مدل پیشنهادی با استفاده از آن‌ها اجرا شدند. مدل پیشنهادی با استفاده از داده‌های پاک‌سازی‌شده یک بازه زمانی چهارماهه اجرا و ارزیابی شد. بازه زمانی بررسی شده داده‌های هجده‌هزار و چهارصد و شصت تراکنش به ثبت رسیده را شامل می‌شود. این داده‌ها در دو بخش جداگانه تفکیک شدند: داده‌های آموزشی^{۵۹} که بازه زمانی ۱ آوریل تا ۱۵ ژوئن ۲۰۰۱ را شامل می‌شود و داده‌های آزمون^{۶۰} که تراکنش‌های بازه زمانی ۱۵ ژوئن تا ۳۱ جولای ۲۰۰۱ را دربرمی‌گیرد. داده‌های آموزشی برای توسعه و اجرای مدل پیشنهادی تحقیق و داده‌های آزمون برای ارزیابی عملکرد مدل پیشنهاددهنده و مقایسه کیفیت خروجی‌های طرح اصلی مدل با سایر طرح‌های پیشنهاددهی جانبی استفاده می‌شود. در مباحثی که در بخش‌های آتی ارائه می‌شود منظور از مشتریان هدف مشتریانی است که حداقل یک تراکنش ثبت شده در هر یک از دوره‌های آموزشی و آزمون داشته‌اند.

اجرای مرحله نخست فاز پیشنهاددهی

در اجرای مرحله نخست، علاوه بر طرح مبتنی بر متدلوژی پیشنهادی مدل، که طرح اصلی ۱ نامیده می‌شود، دو طرح جنبی دیگر، که با تکیه بر رویکردهایی تقریباً مشابه توسعه یافته‌اند، اجرا می‌شود. همان‌گونه که اشاره شد، هدف از اجرای این دو طرح مقایسه عملکرد مدل پیشنهادی تحقیق با عملکرد برخی مدل‌های معمول و مشابه موجود در پیشینه پژوهش است. طرح‌های اجرا شده در این مرحله به طور خلاصه در جدول ۲ می‌آید.

در اجرای طرح اصلی ۱، پس از استخراج و نرمال‌سازی مشخصه‌های بخش‌بندی بر اساس رویکرد *PCB-RFM*، برای تعیین تعداد بهینه خوشه‌ها، فرایند خوشه‌بندی به ازای مقادیر گوناگون پارامتر K تکرار می‌شود و در هر تکرار با محاسبه مقدار شاخص دان

$$P_u(L) = \frac{d_u(N)}{N} \quad (10)$$

همان‌گونه که از رابطه ۹ و ۱۰ استنباط می‌شود، افزایش تعداد اقلام پیشنهادی با کاهش مقدار شاخص دقت و افزایش مقدار شاخص یادآوری همراه است. به منظور ایجاد توازن میان این دو معیار، معیار دیگری به نام F_1 تعریف شده است [۵، ۶، ۲۴، ۴۴]. معیار F_1 به هر یک از دو معیار دقت و یادآوری وزنی یکسان اختصاص می‌دهد. این معیار به ازای هر یک از مشتریان هدف رابطه ۱۳ محاسبه می‌شود.

$$Fl_u(L) = \frac{2 \times R_u(L) \times P_u(L)}{R_u(L) + P_u(L)} \quad (11)$$

هر یک از سه معیار فوق به ازای هر یک از مشتریان هدف محاسبه شدند. در ادامه میانگین این سه معیار در سطح همه مشتریان هدف قابل محاسبه است. مقادیر میانگین این سه معیار در ارزیابی عملکرد سیستم پیشنهاددهنده به کار می‌رود.

اجرای مدل و تحلیل و ارزیابی نتایج

در این بخش متدولوژی پیشنهادی مدل تحقیق، منطبق با فازها و مراحل معرفی شده در بخش ۳، بر داده‌های یک فروشگاه برخط اجرا می‌شود. داده‌های اجرای مدل پیشنهادی تحقیق اطلاعات یک‌ساله مشتریان و محصولات و تراکنش‌های یک فروشگاه برخط عرضه‌کننده محصولات دیجیتالی- کامپیوتری است که به منظور امور تحقیقاتی، به صورت برخط، در دسترس پژوهشگران علاقه‌مند قرار گرفته است. محصولات عرضه شده در این فروشگاه بالغ بر هشتصد قلم کالای گوناگون است. این اقلام در قالب هفت کلاس محصول متمایز دسته‌بندی و تفکیک شدند. به منظور ارزیابی هر چه بهتر مدل پیشنهادی و مقایسه عملکرد آن با برخی رویکردهای مشابه و معمول پیشنهاددهی، در اجرای هر یک از مراحل فاز پیشنهاددهی، علاوه بر طرح اصلی مدل، دو طرح جانبی نیز اجرا می‌شود. معرفی و توضیحات روش اجرای این طرح‌ها، که طرح‌های مقایسه‌ای نامیده می‌شود، به ترتیب در جدول‌های ۲ و ۴ می‌آید.

مینیمم مقدار دو شاخص پشتیبان و اطمینان به ترتیب ۰/۱ و ۰/۵ در نظر گرفته شده است.

ارزیابی و مقایسه شاخص‌های عملکردی طرح‌های اجرایی

کیفیت خروجی‌های مرحله نخست فاز پیشنهاددهی به ازای هر یک از طرح‌های معرفی‌شده در جدول ۲ با استفاده از داده‌های تراکنشی دوره آزمون و با محاسبه شاخص‌های دقت و یادآوری و معیار $F1$ محاسبه شد. نتایج در جدول ۳ می‌آید. نتایج گزارش‌شده در جدول ۳ با در نظر گرفتن پنج گروه محصول ابتدایی لیست پیشنهادی^۶ و با ارزیابی کیفیت پیشنهادهای ارائه‌شده به مشتریان هدف حاصل شده است. همچنین، میانگین شاخص $F1$ طرح‌ها به ازای مقادیر گوناگون پارامتر N محاسبه شد. نتایج در شکل ۳ می‌آید. همان‌گونه که در شکل ۳ مشاهده می‌شود به ازای مقادیر پارامتر N طرح اصلی عملکرد بهتری نسبت به دو طرح مقایسه‌ای دیگر دارد. کیفیت پیشنهادهای طرح اصلی با افزایش طول لیست پیشنهادی تا نقطه $N=5$ افزایش یافته است.

خوشه‌های به دست آمده اعتبارسنجی می‌شوند. بنا بر نتایج، بیشترین مقدار شاخص دان در $K=6$ به دست آمد. بنابراین تعداد بهینه خوشه‌هایی که باید در این طرح برای بخش‌بندی مشتریان در نظر گرفته شود معادل شش خوشه است. در اجرای طرح اصلی ۱-۱، که در آن مشتریان بر اساس رویکرد سنتی $T-RFM$ بخش‌بندی می‌شوند، بهترین مقدار شاخص دان به ازای $K=2$ به دست آمد. از این رو در اجرای این طرح مشتریان در دو خوشه بخش‌بندی شدند.

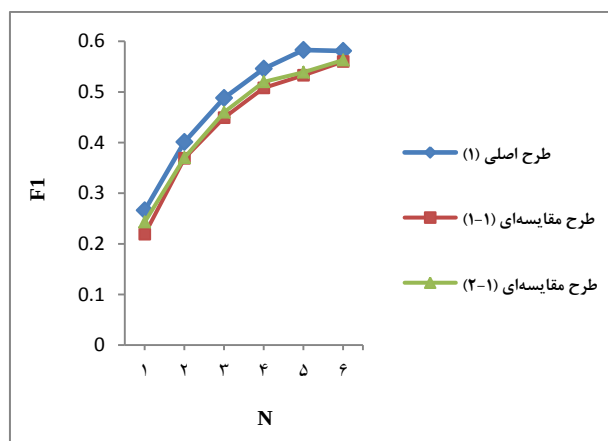
پس از خوشه‌بندی مشتریان، نوبت اجرای نخستین مرحله فاز نخست پیشنهاددهی است. در این فاز روش کشف قوانین انجمنی برای شناسایی الگوهای پنهان خرید مشتریان هر خوشه، در سطح یک رده‌بندی محصول-کلاس محصولات-اجرا می‌شود. در ادامه، با استناد به قوانین معتبری که مینیمم مقدار دو شاخص پشتیبان و اطمینان را تأمین می‌کنند و با تکیه بر رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی، گروه محصولات مطلوب برای هر یک از مشتریان هدف تعیین و پیش‌بینی می‌شود. در اجرای این الگوریتم،

جدول ۲. طرح‌های اجراشده در مرحله نخست فاز پیشنهاددهی

طرح	توضیحات	اصلی‌ترین گام‌های اجرایی
طرح اصلی ۱	این طرح بر اساس متدولوژی پیشنهادی مدل تحقیق، که در بخش پیش تشریح شد، اجرا می‌شود.	بخش‌بندی مشتریان بر اساس رویکرد $PCB-RFM$ / کشف الگوهای پنهان در رفتار خرید مشتریان هر خوشه در سطح یک رده‌بندی محصول و استخراج قوانین پیشنهاددهی معتبر / ارزیابی پیشنهادهای خروجی
طرح مقایسه‌ای ۱-۱	هدف اجرای این طرح مقایسه آثار رویکرد بخش‌بندی $PCB-RFM$ و رویکرد بخش‌بندی $T-RFM$ در کیفیت خروجی‌های مرحله نخست فاز پیشنهاددهی است.	گام‌های اجرایی این طرح مشابه طرح اصلی است؛ با این تفاوت که در آن مشتریان با تکیه بر رویکرد پایه‌ای $T-RFM$ خوشه‌بندی می‌شوند [۵ و ۳۴].
طرح مقایسه‌ای ۱-۲	هدف اجرای این طرح بررسی نقش فاز بخش‌بندی در کیفیت خروجی‌های مرحله نخست فاز پیشنهاددهی است.	در این طرح فاز بخش‌بندی اجرا نمی‌شود. سایر مراحل اجرایی این طرح مشابه دو طرح قبلی است.

جدول ۳. مقایسه شاخص‌های عملکردی طرح‌های اجراشده در مرحله نخست فاز پیشنهاددهی

طرح اصلی ۱			طرح مقایسه‌ای ۱-۱			طرح مقایسه‌ای ۱-۲		
Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
۰,۴۸۲	۰,۸۲۷	۰,۵۸۳	۰,۴۴۵	۰,۷۴۰	۰,۵۳۴	۰,۴۵۴	۰,۷۶۷	۰,۵۳۹



شکل ۳. مقایسه عملکرد طرح‌های اجراشده در مرحله نخست فاز پیشنهاددهی به ازای مقادیر گوناگون پارامتر N

عملکرد مدل‌های پیشنهاددهنده چندان مؤثر واقع شود. این در حالی است که به‌کارگیری رویکرد $PCB-RFM$ برای بخش‌بندی مشتریان، که گونه توسعه‌یافته‌ای از رویکرد سنتی $T-RFM$ تلقی می‌شود، با لحاظ کردن نسبی ترجیحات آنان در جریان اجرای فرایند خوشه‌بندی توانسته در بهبود سطح عملکرد مدل پیشنهاددهنده مؤثر واقع شود.

جمع‌بندی نتایج اجرای مرحله نخست فاز پیشنهاددهی حاکی از آن است که اجرای سیستم پیشنهاددهنده بر اساس رویکرد طراحی شده در طرح اصلی ۱، که همان طرح پیشنهادی مدل تحقیق است، در قیاس با دو طرح مشابه دیگر مطلوب‌ترین خروجی‌ها را تولید می‌کند.

اجرای مرحله دوم فاز پیشنهاددهی

در اجرای این مرحله نیز، علاوه بر طرح مبتنی بر متدلوژی پیشنهادی مدل، که طرح اصلی ۲ نامیده می‌شود، دو طرح جنبی دیگر که با تکیه بر رویکردهایی تقریباً مشابه توسعه یافته‌اند هم اجرا خواهد شد. هدف از اجرای این دو طرح مقایسه عملکرد مدل پیشنهادی تحقیق با عملکرد برخی مدل‌های مشابه موجود در پیشینه پژوهش، از جمله مدل‌های مبتنی بر رویکرد سنتی فیلترینگ مشارکت‌محور، است. طرح‌های اجراشده در این مرحله به طور خلاصه در جدول ۴ می‌آید.

به عبارت دیگر، بهینه مقدار شاخص $F1$ برای این طرح در نقطه $N=5$ به دست آمده است. به‌علاوه، افزایش تعداد اقلام لیست پیشنهادی عملکرد دو طرح مقایسه‌ای را به طور پیوسته بهبود بخشیده است. از طرفی روندهای مشاهده‌شده در مقادیر شاخص $F1$ دو طرح مقایسه‌ای مبین آن است که به ازای مقادیر گوناگون پارامتر N این دو طرح عملکردی تقریباً مشابه از خود نشان داده‌اند. حصول این نتایج به معنای آن است که طرح بخش‌بندی مبتنی بر رویکرد $T-RFM$ تأثیری بهبوددهنده بر روند اجرای مرحله نخست فاز پیشنهاددهی نداشته است؛ زیرا در حالتی که مطابق طرح ۱-۲ مدل با حذف فاز بخش‌بندی اجرا می‌شود کیفیت پیشنهادهای ارائه‌شده توسط مدل تقریباً یکسان و در مواردی بهتر از حالتی است که در آن مدل با تکیه بر طرح ۱-۱ اجرا شده است. دلیل اصلی عدم کارایی رویکرد بخش‌بندی سنتی $T-RFM$ در عملکرد مدل پیشنهاددهنده آن است که این رویکرد هنگام خوشه‌بندی مشتریان ترجیحات و سلاقی آن‌ها را در نظر نمی‌گیرد و فقط به بخش‌بندی مشتریان بر اساس ارزش کلی مشخصه‌های RFM اکتفا می‌کند. همان‌گونه که آمد، مشتریانی که از لحاظ ارزش طول عمر مشابه‌اند لزوماً ترجیحات خرید یکسانی ندارند. بنابراین، لحاظ کردن صرف متغیرهای RFM اگرچه در اجرای بخش‌بندی مبتنی بر ارزش کارایی بالایی دارد، در مواردی مانند مسئله فعلی نمی‌تواند در کشف الگوهای رایج خرید و استخراج قوانین انجمنی و نهایتاً بهبود

جدول ۴. طرح‌های اجراشده در مرحله دوم فاز پیشنهاددهی

طرح	توضیحات	اصلی‌ترین گام‌های اجرایی
طرح اصلی ۲	این طرح بر اساس متدولوژی پیشنهادی مدل تحقیق، که در بخش پیش تشریح شد، اجرا می‌شود.	استخراج ماتریس‌های کاربر-قلم با استناد به خروجی‌های مرحله نخست فاز پیشنهاددهی/ اجرای رویکرد فیلترینگ مشارکت‌محور و استخراج لیست پیشنهاددهی نهایی/ ارزیابی پیشنهاددهی خروجی
طرح مقایسه‌ای ۱-۲	هدف اجرای این طرح بررسی نقش مرحله نخست فاز پیشنهاددهی در سطح کیفیت پیشنهاددهی نهایی است.	گام‌های اجرایی این طرح مشابه طرح اصلی ۲ است؛ با این تفاوت که در آن هنگام استخراج ماتریس‌های کاربر-قلم خروجی‌های مرحله نخست فاز پیشنهاددهی مدنظر قرار نمی‌گیرد.
طرح مقایسه‌ای ۲-۲	هدف از اجرای این طرح مقایسه عملکرد مدل پیشنهادی تحقیق با عملکرد مدل‌هایی است که با تکیه بر رویکرد پایهای فیلترینگ مشارکت‌محور اجرا می‌شوند.	این طرح بر اساس رویکرد سنتی فیلترینگ مشارکت‌محور اجرا می‌شود و در آن هنگام استخراج ماتریس‌های کاربر-قلم خروجی‌های فاز بخش‌بندی و مرحله نخست فاز پیشنهاددهی مدنظر قرار نمی‌گیرد [۵، ۷، ۱۰].

ارزیابی و مقایسه شاخص‌های عملکردی طرح‌های اجراشده در مرحله دوم فاز پیشنهاددهی

به منظور ارزیابی خروجی‌های به‌دست‌آمده در این مرحله، با بررسی داده‌های تراکنشی مشتریان هدف در دوره آزمون، میانگین شاخص‌های عملکردی دقت و یادآوری و معیار $F1$ محاسبه شد. مقادیر شاخص‌های عملکردی به ازای هر یک از سه طرح اجراشده در مرحله دوم فاز پیشنهاددهی در جدول ۵ می‌آید. مقادیر شاخص‌های عملکردی همه طرح‌ها به ازای مقادیر گوناگون پارامتر N محاسبه شد. در همه محاسبات مقدار پارامتر k برابر ۵۰ بود.

اطلاعات جدول ۵ بیانگر آن است که مقادیر شاخص‌های عملکردی طرح اصلی مسئله در قیاس با دو طرح مقایسه‌ای دیگر مطلوب‌تر است. این موضوع حاکی از برتری نسبی عملکرد رویکرد پیشنهاددهی دومرحله‌ای است که در توسعه و طراحی مدل مسئله مدنظر قرار گرفت. به عبارت دیگر، رویکرد طراحی شده با بهره‌گیری از ساختاری سلسله‌مراتبی، که در آن پیشنهاددهی به ترتیب در دو سطح گوناگون از رده‌بندی محصول انجام می‌شود، توانسته در بهبود سطح کیفیت پیشنهاددهی تولیدی و ارتقای عملکرد کلی مدل پیشنهاددهنده مؤثر باشد. مقایسه دقیق‌تر عملکرد هر یک از طرح‌های اجراشده در این مرحله با مشاهده روند تغییرات شاخص $F1$ آن‌ها در شکل ۴ میسر است.

در جریان اجرای طرح اصلی ۱، با بهره‌گیری از رویکرد بخش‌بندی $PCB-RFM$ ، مشتریان در شش خوشه بخش‌بندی شدند و در ادامه، با اجرای روش کشف قوانین انجمنی، کلاس محصولات پیشنهادی برای هر یک از مشتریان هدف شناسایی شدند. در طرح اصلی ۲، با استناد به خروجی‌های مرحله نخست فاز پیشنهاددهی، لیست نهایی اقلام پیشنهادی برای هر یک از مشتریان هدف استخراج می‌شود. با بررسی داده‌های تراکنشی مشتریان خوشه، لیست نهایی پیشنهادها، که دربردارنده اقلام سطح دوم رده‌بندی محصول است، به ازای هر یک از مشتریان هدف تولید می‌شود. این در حالی است که در طرح‌های جانبی ۱-۲ و ۲-۲ خروجی‌های مرحله نخست فاز پیشنهاددهی مدنظر قرار نمی‌گیرند. در نتیجه فضای محاسباتی اقلام نسبت به طرح اصلی مدل گسترده‌تر است. این مرحله فاز پیشنهاددهی بر اساس رویکرد مبتنی بر فیلترینگ مشارکت‌محور و با استفاده از الگوریتم نزدیک‌ترین همسایگی اجرا می‌شود. در این رویکرد، اقلام پیشنهادی به یک کاربر مفروض با درنظرگرفتن تراکنش‌ها و سبدهای خرید سایر کاربرانی که از منظر ترجیحات بیشترین نزدیکی عقیده را به این کاربر دارند تعیین و انتخاب می‌شوند.

* برای هر یک از مشتریان محاسبات این مرحله با استناد به اطلاعات پنج گروه محصول ابتدایی لیست پیشنهادی ($Top-5$)، که در مرحله نخست تولید شده بود، انجام می‌شود.

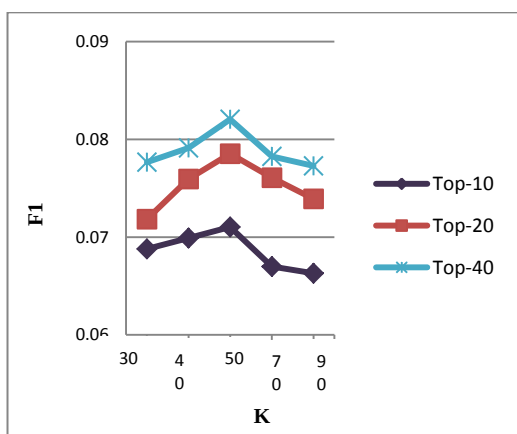
جدول ۵. مقایسه شاخص‌های عملکردی طرح‌های اجراشده در مرحله دوم فاز پیشنهاددهی

Top-10			8-Top			1-Top			
F1	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	
۰,۰۷۱	۰,۲۰۲	۰,۰۴۸	۰,۰۶۷	۰,۱۸۶	۰,۰۴۷	۰,۰۲۶	۰,۰۶۶	۰,۰۶۰	طرح اصلی ۲
۰,۰۵۴	۰,۱۰۴	۰,۰۳۹	۰,۰۵۴	۰,۰۹۶	۰,۰۴۱	۰,۰۱۷	۰,۰۱۱	۰,۰۴۸	طرح مقایسه‌ای ۱-۲
۰,۰۵۴	۰,۱۰۵	۰,۰۳۹	۰,۰۵۲	۰,۰۹۴	۰,۰۳۹	۰,۰۱۱	۰,۰۰۷	۰,۰۳۱	طرح مقایسه‌ای ۲-۲
Top-40			Top-30			Top-20			
F1	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	
۰,۰۸۲	۰,۵۸۴	۰,۰۴۴	۰,۰۸۱	۰,۴۶۰	۰,۰۴۵	۰,۰۷۹	۰,۳۲۴	۰,۰۴۷	طرح اصلی ۲
۰,۰۶۷	۰,۴۱۶	۰,۰۳۷	۰,۰۶۷	۰,۳۱۶	۰,۰۳۷	۰,۰۶۴	۰,۲۰۲	۰,۰۳۸	طرح مقایسه‌ای ۱-۲
۰,۰۶۵	۰,۳۹۰	۰,۰۳۶	۰,۰۶۵	۰,۳۰۱	۰,۰۳۷	۰,۰۶۱	۰,۱۹۴	۰,۰۳۸	طرح مقایسه‌ای ۲-۲

این شاخص، هرچند با آهنگی بسیار ناچیز، به تدریج کاهش یافته یا تقریباً به ثبات رسیده است. پس، افزایش طول لیست پیشنهادی تا آستانه‌ای مشخص می‌تواند در بهبود کیفیت پیشنهادهای ارائه‌شده توسط مدل مؤثر واقع شود. مقدار این آستانه برای طرح اصلی ۲ متناظر با نقطه $N=40$ است.

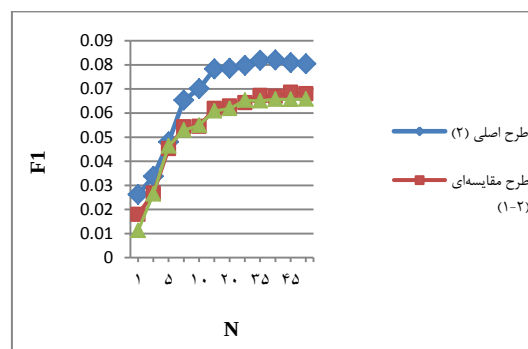
به منظور بررسی نقش تغییرات پارامتر k بر عملکرد اجرایی مرحله دوم فاز پیشنهاددهی، مقدار میانگین شاخص $F1$ به ازای سه سطح گوناگون پارامتر N محاسبه شد. نتایج در شکل ۵ می‌آید. شکل ۵ با توجه به خروجی‌های طرح اصلی ۲ رسم شده است.

روندهای شکل ۵ حاکی از آن است که به ازای هر سه مقدار پارامتر N افزایش پارامتر k تا نقطه $k=50$ موجب بهبود عملکرد مدل می‌شود. با وجود این، در ادامه، افزایش مقدار این پارامتر تأثیری نامطلوب بر کیفیت پیشنهادهای تولیدی داشته و موجب کاهش مقادیر شاخص $F1$ شده است.



شکل ۵. تأثیر تغییرات مقادیر پارامتر k بر شاخص‌های عملکردی مدل در اجرای مرحله دوم فاز پیشنهاددهی

روندهای شکل ۴ مؤید آن است که به ازای همه مقادیر پارامتر N رویکرد پیشنهاددهی دومرحله‌ای یا همان طرح اصلی مدل عملکردی مشابه یا بهتر از دو طرح مقایسه‌ای دیگر دارد. در واقع، به استثنای نقطه متناظر با $N=5$ ، که در آن هر سه طرح عملکردی تقریباً مشابه داشته‌اند، در سایر نقاط پیشنهادهای خروجی در طرح اصلی ۲ کیفیتی بهتر داشته است.



شکل ۴. مقایسه عملکرد طرح‌های اجراشده در مرحله نخست فاز پیشنهاددهی به ازای مقادیر گوناگون پارامتر N

پارامتر k یکی از ورودی‌های الگوریتم نزدیک‌ترین همسایگی است. تعداد همسایگان انتخاب‌شده برای هر یک از مشتریان هدف متناظر با مقدار تعیین‌شده برای این پارامتر است. تحلیل نقش پارامتر N در کیفیت خروجی‌های این مرحله با بررسی روندهای شکل ۴ میسر است. با توجه به روند نمودار $F1$ ، طرح‌های اجرایی عملکرد کلی مدل پیشنهاددهنده با افزایش تعداد اقلام لیست پیشنهادی تا نقطه مشخص بهبود یافته است. اما در ادامه هیچ‌گونه افزایشی در روند تغییرات شاخص $F1$ مشاهده نشده و میانگین مقادیر

جمع‌بندی، نتیجه‌گیری، و تحقیقات آتی

در این تحقیق مدلی جدید برای طراحی و توسعه یک سیستم پیشنهاددهنده محصول به مشتریان خرده‌فروشی‌های برخط ارائه شد. این سیستم با بهره‌گیری از رویکرد بخش‌بندی مشتریان، که در آن از مدل پایه‌ای RFM برای خوشه‌بندی مشتریانی با ترجیحات نسبی مشابه استفاده می‌شود، و با اجرای سلسله‌مراتبی رویکردهای کشف قوانین انجمنی و فیلترینگ مشارکت‌محور امکان ارائه پیشنهاد به مشتریان در دو سطح گوناگون رده‌بندی محصول را فراهم آورد. نتایج اجرای مدل پیشنهادی حاکی از آن است که بهره‌گیری از رویکرد بخش‌بندی بازار در کنار اجرای یک ساختار سلسله‌مراتبی برای ارائه پیشنهاد در دو سطح گوناگون رده‌بندی محصول - ساختاری که امکان بهره‌گیری هم‌زمان از ویژگی‌های رویکردهای پیشنهاددهی مبتنی بر کشف قوانین انجمنی و فیلترینگ مشارکت‌محور را مهیا کرده - سبب شده مدل پیشنهادی این تحقیق در قیاس با مدل‌های معمولی، که صرفاً با تکیه بر رویکرد سنتی فیلترینگ مشارکت‌محور توسعه می‌یابند، عملکرد بهتری داشته باشد. رویکرد پیشنهاددهی دومرحله‌ای، با کاهش فضای محاسباتی اقلام و کاربران و به تبع آن کاهش معضل عدم تراکم و مقیاس‌پذیری داده‌های ورودی، دقت و سرعت الگوریتم نزدیک‌ترین همسایگی را ارتقا داده و عملکرد کلی مدل پیشنهاددهنده را بهبود بخشیده است. به علاوه، رویکرد بخش‌بندی پیشنهادی تحقیق، که به اختصار PCB-RFM نامیده شد، با تکیه بر تعاریف پایه‌ای مدل RFM و با لحاظ کردن نسبی ترجیحات و سلاقی مشتریان در جریان اجرای گام خوشه‌بندی فرایند کشف الگوهای پنهان در تراکنش‌های خرید مشتریان هر خوشه را تسهیل کرد و توانست در بهبود عملکرد مدل پیشنهاددهنده مؤثر واقع شود. نتایج و دستاوردهای محاسباتی حاصل از اجرای مدل پیشنهادی به طور خلاصه به شرح زیر است:

❖ به‌کارگیری رویکرد بخش‌بندی در توسعه سیستم‌های پیشنهاددهنده در صورتی توجیه‌پذیر است که نتایج حاصل از بخش‌بندی بتواند در

ارتقای عملکرد مدل پیشنهاددهنده مؤثر واقع شود. به عبارت دیگر، بخش‌بندی باید این کارایی را داشته باشد که با تشکیل گروه‌هایی همگن از مشتریان فرایند شناسایی و استخراج ترجیحات را تسهیل و در نهایت بهبود کیفیت پیشنهادهای ارائه‌شده را تضمین کند. در این تحقیق، نقش بخش‌بندی مبتنی بر ارزش در عملکرد سیستم‌های پیشنهاددهنده بررسی شد. نتایج حاکی از آن است که خوشه‌بندی کاربران با تکیه بر رویکرد معمول بخش‌بندی مبتنی بر ارزش تأثیر بهبوددهنده‌ای در عملکرد مدل پیشنهاددهنده ندارد. منظور از رویکرد معمول همان رویکردی است که در این تحقیق به اختصار T-RFM نامیده شد و در پیشینه پژوهش به دفعات برای بخش‌بندی کاربران سیستم‌های پیشنهاددهنده استفاده شده است. یکی از دلایل مهم عدم کارایی رویکرد بخش‌بندی T-RFM در عملکرد مدل پیشنهاددهنده آن است که این رویکرد هنگام خوشه‌بندی مشتریان ترجیحات و سلاقی آن‌ها را در نظر نمی‌گیرد و به بخش‌بندی مشتریان بر اساس ارزش کلی مشخصه‌های RFM اکتفا می‌کند؛ این در حالی است که مشتریانی که از لحاظ ارزش طول عمر مشابه‌اند لزوماً ترجیحات خرید یکسانی ندارند. از طرف دیگر نتایج به‌دست‌آمده حاکی از آن است که رویکرد بخش‌بندی پیشنهادشده در این تحقیق، که به اختصار PCB-RFM نام گرفت، با ارائه تعریفی توسعه‌یافته از مشخصه‌های رویکرد معمول T-RFM و با لحاظ کردن نسبی ترجیحات مشتریان در جریان اجرای فرایند خوشه‌بندی، کشف و استخراج قوانین انجمنی در سطح خوشه‌ها را آسان می‌کند و به بهبود عملکرد مدل پیشنهاددهنده منجر می‌شود. با جمع‌بندی مباحث مطرح‌شده می‌توان گفت به‌کارگیری رویکردهای بخش‌بندی بازار، در صورت انتخاب صحیح مشخصه‌های بخش‌بندی، به طور قطع در عملکرد

مدل‌های پیشنهاددهنده اثری مطلوب و بهبوددهنده به جای می‌گذارد.

❖ در مدل‌های پیشنهاددهنده‌ای که با تکیه بر رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی توسعه می‌یابند انتخاب سطحی مناسب از رده‌بندی محصول برای اجرای این رویکرد در اعتبار و کیفیت قوانین استخراج‌شده نقشی اثرگذار و تعیین‌کننده دارد. طبق نتایج حاصل‌شده از اجرای مقدماتی طرح تحقیق، در مواردی که طیف اقلام عرضه‌شده تنوع و گستردگی بالایی دارد اجرای رویکرد پیشنهاددهی مبتنی بر کشف قوانین انجمنی در سطوح پایین رده‌بندی محصول کارایی مناسبی ندارد؛ زیرا به دلیل بالا بودن تعداد اقلام گوناگون محصول اقلام موجود در سبدهای خرید فعال‌ترین مشتریان نیز درصدی ناچیز از کل اقلام عرضه‌شده را شامل می‌شود. این پدیده معضل عدم تراکم ماتریس‌های کاربر-قلم را تشدید می‌کند. مجموعه این عوامل احتمال همپوشانی الگوهای خرید مشتریان را کاهش می‌دهد و استخراج قوانین انجمنی معتبر را در سطوح پایین رده‌بندی محصول با محدودیت مواجه می‌کند. در چنین شرایطی، اتخاذ رویکردهایی مشابه آنچه در طراحی مدل پیشنهادی این تحقیق مدنظر قرار گرفت، یعنی اجرای تکنیک کشف قوانین انجمنی، نه در سطح اقلام نهایی محصول، بلکه در سطح گروه محصولات، می‌تواند با کاهش معضل عدم تراکم داده‌های بررسی‌شده زمینه استخراج قوانین پیشنهاددهی معتبر و قابل استناد را مهیا کند و در نهایت به اکتساب دانشی سودمند درباره رفتار، ترجیحات، و الگوهای خرید مشتریان بینجامد.

❖ جمع‌بندی نتایج اجرای مدل پیشنهادی حاکی از آن است که بهره‌گیری از رویکرد بخش‌بندی بازار در کنار اجرای ساختاری سلسله‌مراتبی برای ارائه پیشنهادها در دو سطح گوناگون از رده‌بندی محصول-ساختاری که امکان بهره‌گیری هم‌زمان از ویژگی‌های رویکردهای پیشنهاددهی مبتنی بر

کشف قوانین انجمنی و فیلترینگ مشارکت‌محور را مهیا می‌کند- سبب می‌شود مدل پیشنهادی این تحقیق در قیاس با مدل‌های معمولی که صرفاً با تکیه بر رویکرد سنتی فیلترینگ مشارکت‌محور توسعه می‌یابند عملکرد بهتری داشته باشد [۵، ۲۴، ۳۵]. رویکرد پیشنهاددهی دومرحله‌ای، با کاهش فضای محاسباتی اقلام و کاربران و به تبع آن کاهش معضل عدم تراکم ماتریس‌های ورودی، دقت و سرعت الگوریتم نزدیک‌ترین همسایگی را ارتقا می‌دهد و عملکرد کلی مدل پیشنهاددهنده را بهبود می‌بخشد.

از منظر دستاوردهای کاربردی و در جهت تحقق برنامه‌های بازاریابی تک‌به‌تک، سیستم پیشنهادی این تحقیق می‌تواند ابزاری کارآمد در بستر وبگاه‌های ایرانی فعال در حوزه تجارت الکترونیکی باشد. این سیستم قادر است با ارائه خدمات شخصی به کاربران آنان را در جست‌وجو و انتخاب محصولات و نیازمندی‌هایشان یاری کند و بدین ترتیب اشتیاق کاربران را به بهره‌گیری از خدمات تجارت الکترونیکی و خرید برخط افزایش دهد. به علاوه، با توجه به اینکه خروجی‌های فاز پیشنهاددهی امکان شناسایی کلاس محصولات پیشنهادی به هر یک از مشتریان هدف را فراهم می‌آورد، می‌توان از این خروجی‌ها به نحوی کارآمد در جلب رضایت هر چه بیشتر این مشتریان بهره گرفت. مثلاً، ارائه سرویس‌های شخصی دوره‌ای، مانند ارسال ایمیل، به منظور معرفی جدیدترین اقلامی که به لیست کالاهای کلاس محصولات پیشنهادی به هر یک از مشتریان هدف اضافه شده‌اند، از اقدامات تشویقی است که می‌تواند در ارتقای سطح وفاداری و رضایتمندی مشتریان فروشگاه‌های برخط مؤثر باشد.

برخی از مهم‌ترین موضوعاتی که در جهت بسط و توسعه هر چه بیشتر مدل پیشنهادی این مسئله می‌توانند در دستور کار تحقیقات آتی قرار گیرند بدین شرح‌اند: ۱. در توسعه مدل پیشنهادی این تحقیق بهره‌گیری از مزایای رویکرد فیلترینگ محتوای‌محور مدنظر قرار نگرفت. توسعه هر چه بیشتر مدل با طراحی یک متدولوژی تلفیقی برای به کارگیری هم‌زمان

رویکردهای فیلترینگ محتوای محور و فیلترینگ مشارکت محور از محورهای مطالعاتی جذاب برای تحقیقات آتی است؛ ۲. خوشه‌بندی مشتریان بر اساس سایر رویکردهای رایج در پیشینه بخش‌بندی، مانند خوشه‌بندی مشتریان بر اساس ویژگی‌های جمعیت‌شناختی یا مشخصه‌های رفتاری و هدایتگر، و بررسی تأثیرات این رویکردها بر عملکرد مدل پیشنهادی این تحقیق می‌تواند از دیگر محورهای مطالعاتی برای تحقیقات آتی باشد؛ ۳. مدنظر قراردادن الگوهای مبتنی بر توالی خرید مشتریان در جریان طراحی مدل پیشنهاددهنده از دیگر ملاحظات است که می‌توان در پژوهش‌های آتی بدان پرداخت.

مراجع

- 1- E. W. T. Ngai, Li Xiu, Chau, D. C. K. (2009). "Application of data mining techniques in customer relationship management: A literature review and classification." *Expert Systems with Applications*, Vol. 36, 2592–2602.
- 2- T. Jiang and A. Tuzhilin. (2006). "Segmenting customers from population to individuals: Does 1-to-1 keep your customers forever." *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, 1297–1311.
- 3- S. K. Shinde and U. Kulkarni. (2012). "Hybrid personalized recommender system using centering-bunching based clustering algorithm." *Expert Systems with Applications*, Vol. 39, 1381–1387.
- 4- L. M. d. Campos, J. M. Fernández-Luna, J. F. Huete, Miguel A. Rueda-Morales. (2010). "Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks." *International Journal of Approximate Reasoning*, Vol. 51, 785–799.
- 5- D. R. Liu, C. Lai, W.J. Lee. (2009). "A hybrid of sequential rules and collaborative filtering for product recommendation." *Information Sciences*, Vol. 179, 3505–3519.
- 6- L. S. Chen, F.H. Hsu, M.Ch. Chen, Y.Ch Hsu (2008). "Developing recommender systems with the consideration of product profitability for sellers." *Information Sciences*, Vol. 178, 1032–1048.
- 7- G. Adomavicius and A. Tuzhilin. (2005). "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions." *IEEE Transactions On Knowledge And Data Engineering*, Vol. 17, 734–749.
- 8- R. Burke. (2002). "Hybrid recommender systems: survey and experiments." *User Modeling and User-Adapted Interaction*, Vol. 12, 331–370.
- 9- L. Lu, M. Medo, Ch.H Yeung, Y.Ch. Zhanga, Z.K. Zhang, T. Zhoua. (2012). "Recommender Systems." *Physics Reports*.
- 10- D. H. Park, H.K Kim, Y. Choi, J.K Kim. (2012). "A literature review and classification of recommender systems research." *Expert Systems with Applications*, Vol. 39, 10059–10079.
- 11- S. H. Choi, S. Kang, Y.J Jeon. (2006). "Personalized recommendation system based on product specification values." *Expert Systems with Applications*, Vol. 31, 607–616.
- 12- P. Resnick. (1994). "GroupLens: an open architecture for collaborative filtering on netnews." in Proc. conference on computer supported cooperative work, 175–186.
- 13- J. B. Schafer, J.A. Konstan, J. Riedl. (2001) "E-commerce recommendation applications." *Data Mining and Knowledge Discovery*, Vol. 5, 115–153.
- 14- H. F. Wang and C. T. Wu. (2012). "A strategy-oriented operation module for recommender systems in E-commerce." *Computers & Operations Research*, Vol. 39, 1837–1849.
- 15- S. H. Choi and B. S. Ahn. (2011). "Rank order-based recommendation approach for multiple featured products." *Expert Systems with Applications*, Vol. 38, 7081–7087.
- 16- M. Balabanovic and Y. Shoham. (1997). "Fab: Content-based, Collaborative Recommendation." *Communications of the ACM*, Vol. 40, 66–72.

- 17- K. Lang. (1995). "Newsweeder: Learning to filter news," in Proc. 12th International Conference on Machine Learning, 331-339.
- 18- H. F. Wang and C. T. Wu. (2009). "A mathematical model for product selection strategies in a recommender system." *Expert Systems with Applications*, Vol. 36, 7299–7308.
- 19- M. Montaner, B. Lopez and J. L. DE LA ROSA. (2003). "A Taxonomy of Recommender Agents on the Internet." *Artificial Intelligence Review*, Vol. 19, 285-330.
- 20- S. L. Huang. (2011). "Designing utility-based recommender systems for e-commerce: Evaluation of preference-elicitation methods." *Electronic Commerce Research and Applications*, Vol. 10, 398–407.
- 21- Y. F. Wang, Y.L Chuang, M.H Hsu, H.Ch Keh. (2004). "A personalized recommender system for the cosmetic business." *Expert Systems with Applications*, Vol. 26, 427–434.
- 22- W. Yang and J. Dia. (2008). "Discovering cohesive subgroups from social networks for targeted advertising." *Expert Systems with Applications*, Vol. 34, 2029–2038.
- 23- A. Albadvi and M. Shahbazi. (2010). "Integrating rating-based collaborative filtering with customer lifetime value: New product recommendation technique." *Intelligent Data Analysis*, Vol. 14, 143-155.
- 24- B. Sarwar, G. Karypis, J. Konstan, J. Riedl. (2000). "Analysis of recommendation algorithms for e-commerce." in Proc. 2nd ACM Conference on Electronic Commerce, 158–167.
- 25- Y. Y. Shih and D. R. Liu. (2008). "Product recommendation approaches: Collaborative filtering via customer lifetime value and customer demands." *Expert Systems with Applications*, Vol. 35, 350-360.
- 26- J. S. Breese. (1998). "Empirical analysis of predictive algorithms for collaborative filtering." in Proc. 14th Annual Conference on Uncertainty in Artificial Intelligence, 43-52.
- 27- D. C. Li, W.L Dai, W.T Tseng. (2011). "A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business." *Expert Systems with Applications*, Vol. 38, 7186–7191.
- 28- S. M. S. Hosseini, A. Maleki, M.R Gholamian. (2010). "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty." *Expert Systems with Applications*, Vol. 37, 5259–5264.
- 29- S. Y. Kim, T.S Jung, E.H. Suh, H.S. Hwang. (2006). "Customer segmentation and strategy development based on customer lifetime value: A case study." *Expert Systems with Applications*, Vol. 31, 101-107.
- 30- S. S. R. Abidi and J. Ong. (2000). "A Data Mining Strategy for Inductive Data Clustering: A Synergy Between Self-Organising Neural Networks and K-Means Clustering Techniques." in Proc. IEEE TENCON, 568-573.
- 31- C. H. Cheng and Y. S. Chen. (2009). "Classifying the segmentation of customer value via RFM model and RS theory." *Expert Systems with Applications*, Vol. 36, 4176–4184.
- 32- D. Jain and S. S. Singh. (2002). "Customer lifetime value research in marketing: a review and future directions." *Journal Of Interactive Marketing*, Vol. 16, 34-46.
- 33- J. T. Wei, S.Y. Lin, Ch.Ch. Weng, H.H. Wu. (2012). "A case study of applying LRFM model in market segmentation of a children's dental clinic." *Expert Systems with Applications*, Vol. 39, 5529–5533.
- 34- D. R. Liu and Y. Y. Shih. (2005). "Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences." *The Journal of Systems and Software*, Vol. 77, 181-191.
- 35- D. R. Liu and Y. Y. Shih. (2005). "Integrating AHP and data mining for product recommendation based on customer lifetime value." *Information & Management*, Vol. 42, 387-400.
- 36- G. T. S. Ho, W.H. Ip, C.K.M. Lee, W.L. Mou. (2012). "Customer grouping for better resources allocation using GA based clustering technique." *Expert Systems with Applications*, Vol. 39, 1979–1987.

- 37- S. Mitra. (2002). "Data mining in soft computing framework: A survey." *IEEE Transactions on Neural Networks*, Vol. 13, 3-14.
- 38- J. Han and M. Kamber. (2005). "Data Mining Concepts and Techniques", 2nd ed. San Francisco: The Morgan Kaufmann Series in Data Management Systems.
- 39- M. Brun. (2007). "Model-based evaluation of clustering validation measures." *Pattern Recognition*, Vol. 40, 807-824.
- 40- Y. S. Kim and B. J. Yum. (2011). "Recommender system based on click stream data using association rule mining." *Expert Systems with Applications*, Vol. 38, 13320-13327.
- 41- W. Y. Chiang. (2011). "To mine association rules of customer values via a data mining procedure with improved model: An empirical case study." *Expert Systems with Applications*, Vol. 38, 16-17, 1722.
- 42- R. Agrawal, T. Imielinski, A. Swami. (1993). "Mining Association Rules between Sets of Items in Large Databases." in Proc.1993 ACM SIGMOD Conference on Management of data, 207-216.
- 43- Y. H. Cho and J. K. Kim. (2004). "Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce." *Expert Systems with Applications*, Vol. 26, 233-246.
- 44- Y. J. Park and K. N. Chang. (2009). "Individual and group behavior-based customer profile model for personalized product recommendation." *Expert Systems with Applications*, Vol. 36, 1932-1939.

واژه‌های انگلیسی به ترتیب استفاده در متن

- 1- E-commerce
 - 2- On-line
 - 3- One-to-one Marketing
 - 4- Recommender Systems
 - 5- Personalization
 - 6- Data Mining Techniques
 - 7- Value-based Segmentation Model
 - 8- Association Rule Mining
 - 9- Collaborative Filtering
 - 10- Amazon.com
 - 11- GroupLens
 - 12- Ringo
 - 13- NetFlix
 - 14- Fab
 - 15- NewsWeeder
 - 16- Ratings
 - 17- Content based Filtering
 - 18- Hybrid methods
 - 19- Subjective Factors
 - 20- Similarity Measures
 - 21- Peers
 - 22- Cold Start
 - 23- Scalability
 - 24- Sparsity
 - 25- User-Item Matrix
 - 26- Customer Lifetime Value (CLV)
-

- 27- Recency
 - 28- Frequency
 - 29- Monetary
 - 30- Analytic Hierarchy Process
 - 31- Weighted RFM
 - 32- Extended Preference
 - 33- Sequential Rule Method (SR)
 - 34- Product Taxonomy
 - 35- Heterogeneous
 - 36- Homogeneous
 - 37- Partitioning methods
 - 38- Model-based methods
 - 39- Hierarchical methods
 - 40- Density-based methods
 - 41- Criterion function
 - 42- Validity Index
 - 43- Davis-Bouldin Index
 - 44- Ward's method
 - 45- Dunn's based method
 - 46- Association Models
 - 47- Association Rules
 - 48- Frequent Itemsets
 - 49- Support
 - 50- Confidence
 - 51- Similarity-Based
 - 52- Memory-Based
 - 53- Scalability
 - 54- Product Category Based-RFM
 - 55- User-Product Category Matrix
 - 56- User-Item Matrix
 - 57- Precision
 - 58- Recall
 - 59- Train Data
 - 60- Test Data
 - 61- Top-5
-