Ind. Manag. Persp. 2023, 13(2), 99-118; doi: 10.48308/JIMP.13.2.99

Industrial Management Perspective

Journal hompage: https://jimp.sbu.ac.ir/?lang=en

Original Article

A Customer-Centric Approach for Recommending Products: A Case Study of Digikala

Hamid Sharifi Esfahani*, Naser Shahsavari Pour**, Mohammad Hossein Sharifi***, Razieh bahmanyar****

Abstract

As competition among marketing companies and retailers intensifies, segmenting customers and recommending suitable products has become a critical strategy for maintaining a competitive edge. With the rapid growth of online shopping, customers often make purchasing decisions based on their needs and desires. Salespeople play a crucial role in influencing customers, making a product recommendation system essential. Such a system has various applications and can also encourage customers to purchase additional products. In this study, we present a method for recommending products to customers that utilizes the K-means clustering algorithm and the RFM (Recency, Frequency, Monetary) model to segment customers and make personalized product recommendations. To evaluate the performance of the proposed system, we conducted experiments using data collected from Digikala, an online shopping company. The results show that clustering based on the RFM features has better results for cluster number zero, which represents loyal customers. Therefore, to encourage these customers to purchase higher-priced goods, companies can offer special discounts to cluster number zero. Our approach provides a customer-centric solution for increasing sales and customer satisfaction.

Keywords: RFM Model; K-means Clustering Algorithm; Customer Segmentation; Digikala Online Company; Recommender System.

How to Cite: Sharifi Esfahani, Hamid; Shahsavari Pour, Naser; Sharifi, Mohammad Hossein; bahmanyar, Razieh (2023). A Customer-Centric Approach for Recommending Products: A Case Study of Digikala, *Ind. Manag. Persp.*, 13(2), 99-118 (In Persian).

Received: Jul. 09, 2022; Revised: Oct. 20, 2022; Accepted: Dec. 28, 2022; Published Online: Jan. 22, 2023.

* Assistant Professor, Department of Industrial Management, Faculty of Economic and Management, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
Corresponding author. Email: sharifi@vru.ac.ir
** Associate Professor, Department of Industrial Management, Faculty of Economic and Management, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
*** M.A Student, Department of Industrial Management, Faculty of Economic and Management, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
*** M.A Student, Department of Computer Engineering, School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

چشم‌انداز مدیریت صنعتی، ۱۴۰۲، دوره ۱۳، شماره ۲، صص ۹۹ – ۱۱۸، doi: 10.48308/JIMP.13.2.99

چشم‌انداز مدیریت صنعتی

Journal hompage: https://jimp.sbu.ac.ir/?lang=fa

نوع مقاله: پژوهشی

ارائه رویکردی مبتنی بر سابقه خرید مشتریان و توصیه‌ی محصول به مشتری: مورد مطالعه مشتریان دیجی‌کالا

حمید شریفی اصفهانی*، ناصر شهسواری پور**، محمدحسین شریفی***، راضیه بهمن‌یار****

چکیده

امروزه که رقابت بین شرکت‌های بازاریابی و خرده‌فروشی‌ها در بهترین حالت خود قرار دارد، شرکت‌ها سعی می‌کنند با توجه به رویکرد تقسیم‌بندی مشتریان و توصیه محصولات مناسب برای آن‌ها، برتری خود را از دست ندهند. خرید آنلاین در بازار مجازی به سرعت در حال افزایش است و معمولاً مشتریان بر اساس نیازها و تمایلات خود تصمیم به خرید می‌گیرند. فروشندگان نقش مهمی در تأثیرگذاری بر مشتریان دارند به همین دلیل سیستم توصیه محصول بسیار حیاتی است. سیستم توصیه محصول، کاربردهای متنوعی دارد و به منظور ترغیب مشتریان به خرید محصولات دیگر نیز مورد استفاده قرار می‌گیرد. در این پژوهش، یک روش برای توصیه محصول به مشتریان ارائه شده است که با استفاده از الگوریتم خوشه‌بندی K-mean و مدل RFM به تقسیم‌بندی مشتریان و سپس توصیه محصول به آن‌ها می‌پردازد. برای اثبات عملکرد سیستم پیشنهادی، آزمایش‌هایی با استفاده از داده‌های جمع‌آوری‌شده از شرکت «دیجی‌کالا» صورت گرفت. نتایج پژوهش نشان داد که خوشه‌بندی بر اساس ویژگی‌های R، F و M برای خوشه‌ی شماره صفر نتایج بهتری دارد. به همین دلیل، برای ترغیب مشتریان وفادار به خرید کالاهای با قیمت بالاتر، می‌توان از قیمتی که به خوشه شماره صفر (مشتریان وفادار) پیشنهاد می‌شود، بهره برد و تخفیف‌های ویژه‌ای برای این مشتریان در نظر گرفت.

کلیدواژه‌ها: مدل RFM؛ الگوریتم خوشه‌بندی K-means؛ تقسیم‌بندی مشتریان؛ شرکت اینترنتی دیجی‌کالا؛ سیستم توصیه‌گر.

استناددهی: شریفی اصفهانی، حمید؛ شهسواری پور، ناصر؛ شریفی، محمدحسین؛ بهمن‌یار، راضیه (۱۴۰۲). ارائه رویکردی مبتنی بر سابقه خرید مشتریان و توصیه‌ی محصول به مشتری: مورد مطالعه مشتریان دیجی‌کالا. *چشم‌انداز مدیریت صنعتی*، ۱۳(۲)، ۹۹ – ۱۱۸.

تاریخ دریافت: ۱۴۰۱/۰۴/۱۸، تاریخ بازنگری: ۱۴۰۱/۰۷/۲۸، تاریخ پذیرش: ۱۴۰۱/۱۰/۰۷، تاریخ اولین انتشار: ۱۴۰۱/۱۱/۰۲.

* استادیار، گروه مدیریت صنعتی، دانشکده اقتصاد و مدیریت، دانشگاه ولی‌عصر(عج) رفسنجان، رفسنجان، ایران.
نویسنده مسئول: Email: sharifi@vru.ac.ir

** دانشیار، گروه مدیریت صنعتی، دانشکده اقتصاد و مدیریت، دانشگاه ولی‌عصر(عج) رفسنجان، رفسنجان، ایران.

*** دانشجوی کارشناسی ارشد، گروه مدیریت صنعتی، دانشکده اقتصاد و مدیریت، دانشگاه ولی‌عصر(عج) رفسنجان، رفسنجان، ایران.

**** دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

۱. مقدمه

امروزه همه مشتریان برای شرکت‌ها اهمیت یکسانی ندارند و شرکت‌ها به دنبال این هستند که ضمن شناسایی و تحلیل ویژگی‌های مشتریان بتوانند آن‌ها را بر اساس ارزشی که برای شرکت دارند نیز تفکیک و خوشه‌بندی کنند. در دنیای رقابتی امروز، با تغییر محیط کسب‌وکار از محصول‌محوری به مشتری‌محوری، درک رفتار مشتری اهمیت یافته و یکی از چالش‌ها، شناخت مشتریان و درک تفاوت‌های آن‌ها است. مشتری مهم‌ترین شریک تجاری سازمان به‌شمار می‌آید؛ بنابراین تحلیل رفتار مشتری به سازمان کمک می‌کند تا درک بهتری از بازار داشته باشد و بتواند فرصت‌های جدید را کشف کند. با توجه به رقابت فشرده در بازار و وجود گزینه‌های متنوعی از محصولات و خدمات پیش روی مشتریان، شناخت صحیح رفتار مشتریان، بعد از ابعاد مدیریت ارتباط با مشتری، مهم‌ترین عامل است. بررسی و تحلیل رفتار مشتریان و انتخاب شیوه مناسب بازاریابی بر اساس این تحلیل‌ها، عامل بسیار مهمی برای بقای شرکت‌ها تلقی می‌شود [۲۷]. تقسیم‌بندی مشتریان[1] مزیت‌های زیادی برای بخش مدیریت و بازاریابی فراهم می‌کند. مهم‌ترین مزیت این است که امکان شناسایی مؤثر گروه‌های مشتریان کلیدی را فراهم می‌آورد که شامل سودآورترین و وفادارترین مشتریان است؛ همچنین با استفاده از تقسیم‌بندی مشتریان می‌توان در مورد علاقه‌مندی‌های گروه‌های مختلف صحبت کرد و تعیین کرد تناسب محصولاتی یا خدماتی در گروه‌های با سودآوری بالا وجود دارد یا خیر [۵، ۲۹].

خوشه‌بندی مشتریان یکی از مباحث مطرح در حوزه مدیریت ارتباط با مشتری است. خوشه‌بندی درواقع شکستن جمعیت زیاد مشتریان به بخش‌های مختلف است که دیدگاهی کلی و سطح بالا از تمام بانک اطلاعاتی مشتریان ارائه می‌کند و به صاحبان شرکت‌های توزیع امکان اعمال رفتار و سیاست‌های متفاوت و مناسب به مشتریان هر بخش می‌دهد [۲۳]. به‌منظور انجام بخش‌بندی مشتری، تکنیک‌های متعددی در مبانی نظری پیشنهاد شده است و در میان آن‌ها، خوشه‌بندی متداول‌ترین روش مورد استفاده است [۳۱]. علاوه بر این، ویژگی‌های مدل RFM به‌طور مؤثر برای درک و تجزیه‌وتحلیل ویژگی‌های رفتار مشتری استفاده می‌شود [۲۰، ۱۱، ۸، ۱۴، ۲۱].

سیستم‌های توصیه‌گر به‌عنوان یکی از رایج‌ترین راه‌حل‌های نرم‌افزاری در تجارت الکترونیک به‌کار رفته‌اند. این سیستم‌ها با مهیاکردن پیشنهادهایی بر پایه ارجحیت‌ها و خریدهایی که روی سایت‌های تجارت الکترونیک صورت گرفته است به مشتریان در پیداکردن محصولاتی که تمایل به خرید آن‌ها دارند، کمک می‌کنند [۳۰].

1. Customer Segmentation

سؤال‌های مطرح‌شده در پژوهش حاضر به شرح زیر است:
۱. چگونه می‌توان مشتریان «شرکت دیجی‌کالا» را تقسیم‌بندی کرد؟
۲. بر اساس پیشینه خرید مشتریان «دیجی‌کالا»، چند گروه از مشتریان قابل‌شناسایی هستند؟
۳. راهبردهای توصیه محصول برای هر دسته از مشتریان کدام است؟

هدف از انجام پژوهش حاضر، توصیه محصولات مناسب به افرادی است که در گذشته از شرکت خرید داشته‌اند.

۲. مبانی نظری و پیشینه پژوهش

انگیزه پژوهش. با رشد فناوری‌های جدید و شبکه‌های رایانه‌ای، جهان دست‌خوش دگرگونی‌های سریعی در تمام زمینه‌ها شده است. افراد خریدهای خود را به‌صورت اینترنتی انجام می‌دهند، اطلاعات موردنیاز خود را از طریق موتورهای جست‌وجو جمع‌آوری می‌کنند و بخش زیادی از زندگی اجتماعی به اینترنت وابسته شده است. از سوی دیگر حجم عظیم اطلاعات در اینترنت مشکلاتی را به همراه داشته است. برای مثال، کاربران را در فرایند تصمیم‌گیری، انتخاب اطلاعات یا کالای موردنیاز، دچار سردرگمی کرده است. شناخته‌شده‌ترین راه‌حل این مشکل استفاده از سیستم‌های توصیه‌گر است. سیستم‌های توصیه‌گر مشتری را در مواجهه با سیل عظیم اطلاعات در اینترنت، یاری می‌رسانند و او را از سردرگمی در میان میلیون‌ها محصول قابل‌خرید، رهایی می‌بخشند. این سیستم‌ها راهی برای کمک به افراد برای پیداکردن اقلامی است که برای افراد جالب و یا مطابق نیازهای آن‌ها است. در این سیستم‌ها پیشنهاد یک آیتم خاص به یک مشتری خاص بر اساس اطلاعات پروفایل کاربر، اطلاعات پروفایل کالاها و یا رتبه‌های اختصاص‌داده‌شده توسط این کاربر و کاربران دیگر به اقلام و کالاهایی است که قبلاً خریداری کرده و یا استفاده کرده‌اند. در این پژوهش، یک سیستم توصیه‌گر برای بزرگترین شرکت‌های فروش اینترنتی در ایران مورد بررسی قرار می‌گیرد. ابتدا با استفاده از مدل RFM و روش خوشه‌بندی K-Means، مشتریان تقسیم‌بندی می‌شوند؛ سپس یک سیستم توصیه‌گر برای این مجموعه داده، بر اساس آیتم‌ها ارائه می‌شود. این سیستم قابلیت توسعه دارد و برای شرکت‌های دیگری در ایران قابل‌استفاده خواهد بود.

مفاهیم پژوهش

تقسیم‌بندی مشتریان. مشتریان افراد یا مشاغلی هستند که محصولات یا خدمات را از یک شرکت خریداری می‌کنند که استراتژی‌های بازاریابی برای افزایش رضایت و وفاداری مشتری اعمال می‌شود. تقسیم‌بندی مشتری می‌تواند برای تقسیم مشتریان به گروه‌های متمایز کوچک‌تر

انتخاب شود تا استراتژی بازاریابی بتواند هر گروه را به‌صورت جداگانه هدف قرار دهد. تقسیم‌بندی مشتریان به شرکت کمک می‌کند تا مشتریان ارزشمند را هدف قرار داده و فعالیت‌های بازاریابی را توسعه دهد [۷]. تقسیم مشتریان به لحاظ رفتار به‌وجودآمده، عامل اصلی در مدیریت ارتباط با مشتریان و همچنین بازاریابی هدفمند به‌شمار می‌رود که نخستین بار توسط اسمیت[1] (۱۹۵۶) معرفی شد و توسط بسیاری از مطالعات علمی تأیید شده و توسط چندین شرکت با موفقیت به‌کار گرفته شده است [۱۵]. مشتریان نیازها، رفتارها و ترجیحات متفاوتی دارند و برای شرکت‌ها چالش‌برانگیز است که به همه مشتریان به‌طور یکسان خدمت‌رسانی کنند. تقسیم‌بندی مشتری در پاسخ به این مشکل پدیدار شده است.

در تقسیم‌بندی مشتری، مشتریان به لحاظ هر نوع متغیری دسته‌بندی می‌شوند که می‌تواند به‌طور گسترده‌ای به دو دسته تقسیم شود: متغیرهای عمومی و متغیرهای خاص محصول [۱۶]. متغیرهای عمومی شامل جمعیت‌شناسی مشتری (جنسیت، سن، درآمد، سطح تحصیلات و غیره) و شیوه زندگی است؛ درحالی‌که متغیرهای خاص محصول شامل رفتار خرید مشتری (مانند فرکانس خرید، مصرف، هزینه و غیره) و اهداف است. بااین‌حال متغیرهای عمومی ساده‌تر هستند. متغیرهای خاص محصول برای گرفتن رفتار خرید از مشتریان مهم‌تر است و احتمال بیشتری برای تمایز مشارکت مشتریان در یک کسب‌وکار در این زمینه، فراهم می‌کنند.

مدل RFM. مفهوم RFM نخستین بار توسط هوگس[2] (۱۹۹۴) معرفی شد. وی برای تحلیل RFM از رفتار گذشته مشتری که به‌آسانی قابل‌پیگیری و دسترسی است، استفاده کرد. این مدل از سه بُعد مربوط به داده‌های مبادلاتی مشتریان، برای تحلیل رفتار آن‌ها استفاده می‌کند؛ به همین دلیل این مدل از متغیرهای خاص محصول برای تقسیم‌بندی مشتریان بهره می‌گیرد.

شاخص‌های این مدل به‌صورت زیر تعریف می‌شوند: تازگی مبادله[3] (R)، تعداد تکرار مبادله (فراوانی)[4] (F) و ارزش پولی مبادله[5] (M) [۴]. درواقع هدف از RFM، تقسیم‌بندی مشتریان بر اساس رفتار خرید آن‌ها است. این سه متغیر به‌صورت زیر تعریف می‌شوند:

تازگی مبادله یا تأخر (R): مشتری طی مدت گذشته، چه زمانی با شرکت ارتباط برقرار کرده و چقدر از آخرین مراجعه مشتری به شرکت گذشته است (هرچه کمتر باشد، احتمال بازگشت مشتری و تکرار خرید توسط مشتری یاد شده بیشتر است).

1. Smith
2. Hughes
3. Recency
4. Frequency
5. Monetary

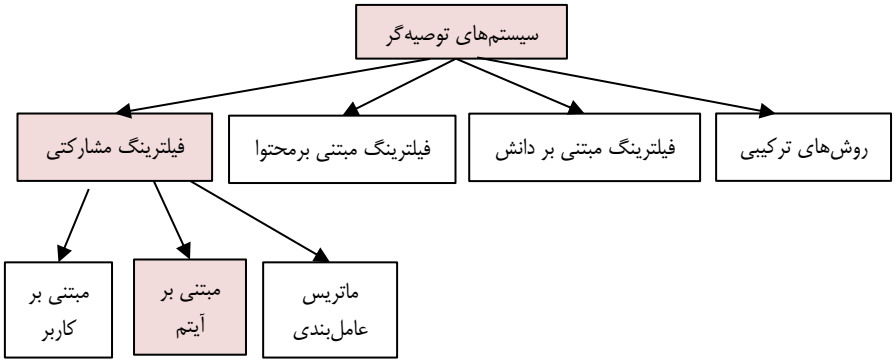
تعداد تکرار مبادله یا فراوانی (F): تعداد دفعات خرید مشتری در یک دوره مشخص چقدر بوده است (هرچه تعداد آن بیشتر باشد، حاکی از آن است که مشتری موردنظر در خرید از شرکت، ثبات‌قدم[1] بیشتری دارد).
ارزش پولی مبادله یا تبادل مالی (M): میزان تبادل مالی مشتری با شرکت در طول یک دوره زمانی مشخص چقدر بوده است (هرچه مقدار این پارامتر بیشتر باشد، نشان‌دهنده این است که مشتری موردنظر، اهمیت بیشتری دارد و باید به آن مشتری توجه بیشتری داشت).

خوشه‌بندی[2]. خوشه‌بندی، فرایند گروه‌بندی مجموعه‌ای از اشیای فیزیکی یا انتزاعی در گروه‌هایی از اشیا مشابه است. هر خوشه مجموعه‌ای از اشیای داده‌ای است که اعضای درون دسته با هم مشابه هستند؛ درحالی‌که با اعضای گروه‌های دیگر تفاوت دارند. فرایند تقسیم‌بندی مجموعه‌ای از داده‌ها با هدف تبدیل به زیربخش‌ها است. هر زیربخش یک خوشه است که داده‌ها در هر خوشه نسبت به خوشه‌های دیگر شباهت بیشتری دارند [۱۰].

الگوریتم K-means یکی از محبوب‌ترین و پرکاربردترین روش‌های خوشه‌بندی است [۱۷، ۱۲] که می‌توان به‌راحتی آن را اجرا کرد و در اجرا بسیار سریع است؛ ازاین‌رو معمولاً برای خوشه‌بندی در زمینه‌های پژوهشی و حوزه‌های کاربردی مختلف ازجمله داده‌کاوی، تجزیه‌وتحلیل داده‌های آماری، تشخیص الگو، تقسیم‌بندی مشتری و سایر برنامه‌های کاربردی تجاری استفاده می‌شود [۶]. این روش به تعیین پیشینی تعداد خوشه‌ها (K) نیاز دارد. برای یک مقدار K معین، الگوریتم K-means با تولید تصادفی K نقطه مرکزی (برای مثال، مرکز) شروع می‌شود. در مرحله بعد، فاصله بین هر نمونه و هر مرکز محاسبه شده و سپس هر نمونه به نزدیک‌ترین مرکز اختصاص می‌یابد. پس از تشکیل خوشه‌ها، مقدار میانگین هر خوشه بر اساس اشیای فعلی در خوشه دوباره محاسبه می‌شود. این فرآیند تا زمانی ادامه می‌یابد که همگرایی[3] رخ دهد؛ یعنی مرکزها تغییر نمی‌کنند. برای تعیین یک مقدار K بهینه باید روش K-means را با تعداد خوشه‌های مختلف (K) اجرا کرد؛ زیرا عملکرد این الگوریتم به مقدار K بستگی دارد [۱۸].

سیستم توصیه‌گر. سیستم توصیه‌گر به‌عنوان سیستمی تعریف می‌شود که پس از آگاهی از تمایل مشتریان، محصول یا خدمتی را که با علایق کاربر متناسب است، توصیه می‌کند [۲۲]. انواع سیستم‌های توصیه‌گر در شکل ۱، قابل‌مشاهده است.

1. Loyalty
2. Clustring
3. Convergence



شکل ۱. انواع سیستم‌های توصیه‌گر

فیلترینگ مشارکتی[1]. در این روش بر اساس شباهت رفتاری و الگوهای عملکردی کاربرانی که شباهت‌های رفتاری و الگوهای مشابهی با کاربر فعلی در گذشته داشته‌اند، پیشنهادهای ارائه می‌شود. شاید تعریف آن کمی پیچیده باشد، ولی به‌طور ساده روش فیلترینگ مشارکتی بر این فرض استوار است که کاربرانی با یک سری نظرهای مشابه درباره یک آیتم (منظور از آیتم، فیلم، عکس، موزیک یا هر چیز دیگری است که توصیه می‌شود)، درباره آیتم‌های دیگر نیز نظرهای مشابه خواهند داشت [۹]. فیلترینگ مشارکتی خود شامل سه بخش است: ۱. بر اساس آیتم[2] [۲۲، ۲۸]؛ ۲. بر اساس کاربر[3]، به‌ویژه بر اساس مشخصات مشتری [۳، ۲۴]؛ ۳. بر اساس ماتریس عامل‌بندی[4] [۲].

در پژوهش موردنظر از یک سیستم توصیه‌گر مبتنی بر آیتم استفاده می‌شود. این سیستم به طور خودکار پس از مشاهده یک آیتم توسط کاربر، به دنبال آیتم‌های مرتبط با آن می‌گردد؛ سپس سیستم تحلیل می‌کند کاربرانی که پیش‌تر آن آیتم را مشاهده کرده‌اند، بعداً چه آیتم‌های دیگری را نیز مشاهده کرده‌اند و درنهایت آن آیتم‌ها را به کاربر پیشنهاد می‌دهد. این روش به کاربران امکان می‌دهد تا با دیدن یک آیتم، به‌طور خودکار به آیتم‌های مرتبط و جذاب دیگر هدایت شوند که می‌تواند بهبود تجربه خرید و کشف محصولات جدید را فراهم کند. برای مثال، در «سایت آمازون» اگر یک تلفن همراه تهیه کنید، بعد از آن، سیستم توصیه‌گر قاب گوشی و محافظ صفحه نمایشگر و مواردی مرتبط با تلفن همراه توصیه می‌کند.

1. Collaborative Filtering
2. Item-Based
3. User-Based
4. Matrix Factorization

روش‌های مورداستفاده در سیستم‌های توصیه‌گر. با توجه به بررسی‌های صورت‌گرفته پرکاربردترین روش‌ها در پیاده‌سازی سیستم‌های توصیه‌گر به‌صورت زیر است:

۱. خوشه‌بندی

۲. قواعد انجمنی

خوشه‌بندی: گروه‌بندی نمونه‌های مشابه با یکدیگر در یک نمونه داده‌ای بزرگ و توزیع داده‌ها به K گروه که نقاط هر گروه به یکدیگر شبیه هستند. هدف از خوشه‌بندی این است که داده‌های موجود به چند گروه تقسیم شوند و در این تقسیم‌بندی داده‌های گروه‌های مختلف باید حداکثر تفاوت را با یکدیگر داشته و داده‌های موجود در یک گروه باید بسیار به یکدیگر شبیه باشند. برای استفاده از الگوریتم خوشه‌بندی باید تعداد خوشه بهینه را نیز محاسبه کرد. روش‌های متفاوتی برای محاسبه وجود دارد. در این پژوهش از شاخص Elbow برای محاسبه تعداد بهینه خوشه استفاده شده است. ایده اصلی روش‌های خوشه‌بندی مبتنی بر پارتیشن، مانند K-Means، به‌دست‌آوردن تعداد خوشه‌ها به‌گونه‌ای است که مجموع فواصل درون خوشه‌ای داده‌ها (یا مجموع مربع‌ها) از فواصل درون خوشه‌ای به حداقل می‌رسد. مجموع فواصل درون‌خوشه‌ای داده‌ها درجه خوشه‌بندی انجام‌شده را نشان می‌دهد. روش Elbow مجموع فواصل درون‌خوشه‌ای داده‌ها را تابعی از تعداد خوشه‌ها در نظر می‌گیرد. به این ترتیب تعداد خوشه‌ها به‌گونه‌ای انتخاب می‌شود که با افزودن یک خوشه دیگر، به‌حداقل‌رساندن WSS بهبود نمی‌یابد. تعداد بهینه خوشه‌ها طبق الگوریتم زیر به‌دست می‌آید:

۱. اجرای الگوریتم خوشه‌بندی مانند K-Means برای مقادیر متفاوت K (برای مثال، با درنظرگرفتن مقدار K در بازه ۱ تا ۱۰)؛

۲. محاسبه مقدار WSS برای هر مقدار K

۳. رسم مقدار WSS برحسب مقادیر مختلف K

۴. نقطه زانویی نمودار رسم‌شده، تعداد بهینه خوشه‌ها را نشان می‌دهد.

قواعد انجمنی[1]. الگوریتم‌های قواعد انجمنی به‌صورت عمده در پیداکردن روابط بین اقلام و ویژگی‌هایی که در پایگاه داده‌ها موجود است، استفاده می‌شود تا توانایی درک نقش‌هایی که احتمالاً در طول خرید در فروشگاه‌ها رخ می‌دهد را داشته باشند.

1. Association Rule

۳. روش‌شناسی پژوهش

پژوهش‌های علمی از نظر هدف به دو گروه بنیادی و کاربردی تقسیم می‌شوند. پژوهش‌های بنیادی در جست‌وجوی کشف حقایق و واقعیت‌ها و شناخت پدیده‌ها و اشیا و روابط بین آن‌ها در جهان هستند؛ درحالی‌که پژوهش‌های کاربردی با استفاده از زمینه و بستر شناختی و معلوماتی که از طریق پژوهش‌های بنیادی فراهم شده است، برای رفع نیازمندی‌های بشر مورد استفاده قرار می‌گیرد؛ ازاین‌رو پژوهش حاضر از نظر هدف، کاربردی است.

بر اساس ماهیت و روش، پژوهش‌هایی که به دنبال چیستی و چگونه‌بودن موضوع، بررسی وضع موجود و توصیف منظم و نظام‌دار وضعیت فعلی هستند، توصیفی قلمداد می‌شوند و اگر در این نوع از پژوهش‌ها، ویژگی‌ها و صفات افراد جامعه با هدف شناخت ویژگی‌ها، عقاید، نگرش‌ها و رفتارهای افراد مطالعه شود، پیمایشی محسوب می‌شوند؛ ازاین‌رو پژوهش حاضر توصیفی ـ پیمایشی محسوب می‌شود و از نوع پژوهش‌های پس‌رویدادی است. در این روش، پژوهشگر گذشته‌نگر است و سعی دارد که از معلول به علت احتمالی پی ببرد.

به این دلیل که بیشتر روش‌های حل مسئله در زمینه‌های علوم رفتاری به دو بخش، هدف و نحوه گردآوری داده تقسیم می‌شود، در این پژوهش نحوه گردآوری داده از مراجع معتبر و تحلیل آن از طریق علم داده‌کاوی خواهد بود. تمامی داده‌های این پژوهش به‌صورت کمّی است که از سوابق خرید مشتریان برای مدت ۹ روز از پایگاه داده[1] «شرکت دیجی‌کالا» استخراج شده است. در ادامه، مراحل روش پیشنهادی به‌صورت شکل ۲، مشاهده می‌شود.

1. Database

شکل ۲. مراحل تکنیک پیشنهادی

۴. تحلیل داده‌ها و یافته‌های پژوهش

توصیف داده‌ها. مجموعه داده معتبر از شرکت اینترنتی «دیجی‌کالا» دریافت شد. داده‌ها شامل اطلاعات خریدهای گذشته مشتریان است که شامل ۲۰۰۰۰۰ ردیف از مشتریان و ۵۶۱۰۴ مشتری متفاوت است. ویژگی‌های مجموعه داده در جدول ۱، توصیف شده است.

جدول ۱. توصیف ویژگی‌های مجموعه داده

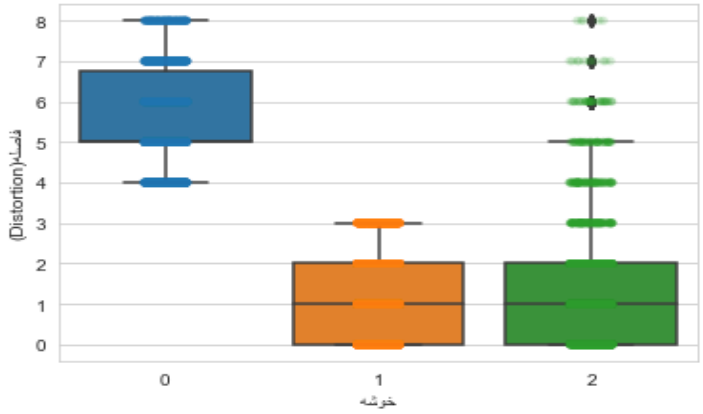
ستون	نام	توضیحات
Order_id	شناسه سفارش‌ها	مشخص‌کننده سبد خرید هر کاربر
User_id	شناسه کاربر	شناسه یکتایی که به هر کاربر نسبت داده شده است.
Variant_id	شناسه دسته‌بندی	شناسه یکتایی برای دسته‌بندی‌های مختلف کالا
Product_id	شناسه کالا	شناسه یکتایی که به هر کالا نسبت داده شده است
quntity	تعداد	تعداد کالای خریداری‌شده در هر سفارش
price	قیمت (ریال)	قیمت کالا
date	تاریخ	تاریخ سفارش
City_id	شناسه شهر	مشخص‌کننده اینکه هر کاربر از کدام شهر بوده است.
City_name	نام شهر	اسم هر شهر
State_id	شناسه استان	مشخص‌کننده اینکه کاربر از کدام استان بوده است.
State_name	نام استان	نام هر استان

در ادامه مراحل انجام پژوهش و یافته‌ها ارائه شده است:

مرحله اول: برای مجموعه داده موردنظر، مدل RFM بررسی و محاسبه می‌شود؛ سپس داده‌های دورافتاده[1] از مجموعه داده حذف می‌شود تا تحلیل مناسب‌تری انجام شود. هدف این پژوهش، پیشنهاد دادن یا توصیه محصولات مناسب به افرادی است که در گذشته از شرکت خرید داشته‌اند. به همین دلیل تنها افرادی انتخاب می‌شوند که تکرار خرید آن‌ها بیشتر مساوی دو باشد. در نهایت تعداد کاربران نهایی موجود برابر ۹۹۲۸ به‌دست آمد.

مرحله دوم: با توجه به متفاوت‌بودن مقیاس، باید داده‌ها استانداردسازی شود؛ بنابراین قبل از خوشه‌بندی، متغیرهای RFM با استفاده از پرکاربردترین تکنیک مقیاس‌بندی، استانداردسازی می‌شوند که هر متغیر با میانگین[2] صفر و انحراف استاندارد[3] یک، تغییر مقیاس می‌دهد.

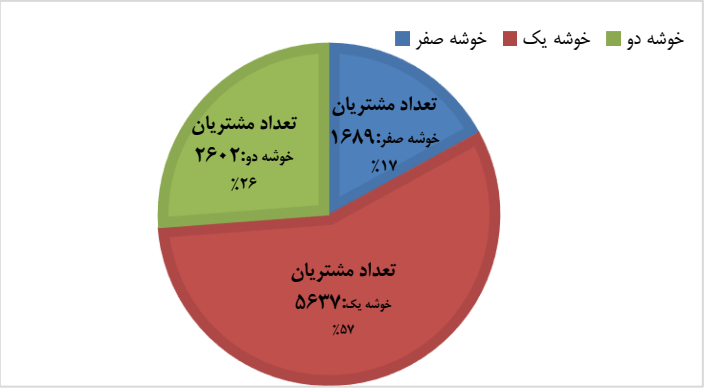
مرحله سوم: در این مرحله خوشه‌بندی با استفاده از داده‌هایی که از تحلیل RFM در مرحله قبل به دست آمده است، صورت می‌گیرد. ابتدا مجموعه داده با استفاده از الگوریتم خوشه‌بندی K-Means به خوشه‌های مختلف تقسیم‌بندی می‌شود. برای محاسبه تعداد خوشه بهینه از الگوریتم تعداد خوشه بهینه‌ی Elbow استفاده شده است. همان‌طور که در مبانی نظری پژوهش بیان شد، روش Elbow، مجموع فواصل درون خوشه‌ای داده‌ها را به‌عنوان تابعی از تعداد خوشه‌ها در نظر می‌گیرد که با توجه به نتایج تعداد بهینه خوشه برابر سه در نظر گرفته شده است. نمودار Elbow برای تعداد بهینه خوشه در شکل ۳، مشاهده می‌شود.



شکل ۳. نمودار تعداد بهینه خوشه با روش Elbow

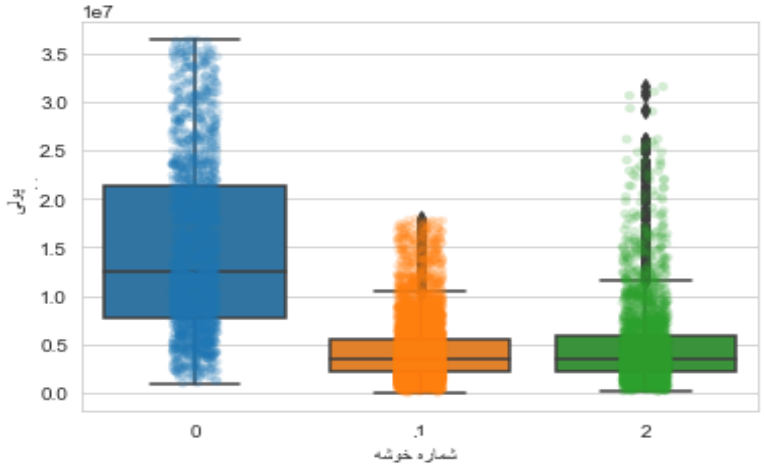
تعداد مشتریان در هر خوشه به‌صورت نمودار در شکل ۴، ارائه شده است.

1Outlier Data
2. Mean
3. Standard Daviation

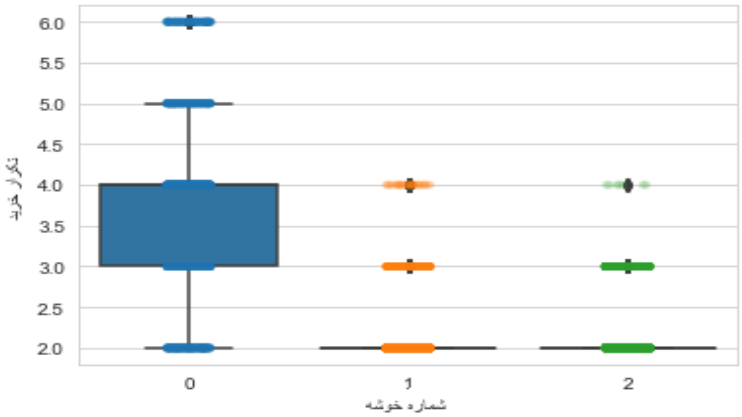


شکل ۴. تعداد مشتریان در هر خوشه

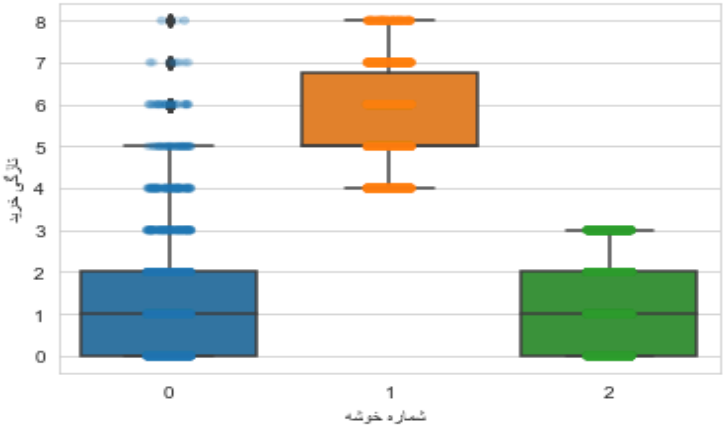
خوشه شماره صفر از نظر ویژگی پولی از سایر خوشه‌ها بیشتر است؛ به این معنا که مشتریان خوشه شماره صفر، میانگین خریدهای بالاتری نسبت به سایر خوشه‌ها داشته‌اند. نتیجه خوشه‌بندی بر اساس ویژگی‌های پولی، فراوانی و تازگی در شکل‌های ۵، ۶ و ۷، نشان داده شده است.



شکل ۵. نمودار خوشه‌بندی بر اساس ویژگی ارزش پولی



شکل ۶. نمودار خوشه‌بندی بر اساس ویژگی تکرار خرید



شکل ۷. نمودار خوشه‌بندی بر اساس ویژگی تازگی خرید

میانگین ویژگی‌های R، F و M برای خوشه‌ها در جدول ۲، مشاهده می‌شود.

جدول ۲. میانگین ویژگی‌های ارزش پولی، تازگی خرید و تکرار خرید

خوشه	تازگی خرید (Regency)	تکرار خرید (Frequency)	پولی (Monetary)
۰	۱/۴۸	۳/۷۷	۱۴۸۵۸۴۲۰
۱	۵/۵۵	۲/۱۴	۴۷۸۱۰۸۰
۲	۱/۲۳	۲/۱۹	۴۳۱۸۵۹۷

برای مثال، خوشه شماره صفر شامل افرادی است که به‌طور متوسط در بازه‌ی زمانی ۹ روزه، میانگین ۳/۷۷ خرید و ۱۴۸۵۸۴۲۰ ریال هزینه و تازگی خرید ۱/۴۸ داشته‌اند.

به دلیل اینکه هدف نهایی هر کسب‌وکاری اغلب دستیابی به سوددهی مطلوب است، ویژگی پولی نسبت به سایر ویژگی‌ها از اهمیت بالاتری برخوردار است. با توجه به ویژگی‌های R، F و M خوشه صفر، شامل افرادی است که نسبت به سایر خوشه‌ها از نظر قیمت کالاهای خریداری‌شده، تکرار خرید و تازگی خرید در رتبه نخست قرار دارند و می‌توان نتیجه‌گیری کرد مشتریان این خوشه، افراد وفادار در «شرکت دیجی‌کالا» هستند. پس از آن خوشه‌ی شماره دو و در نهایت خوشه شماره یک قرار دارد. همان‌طور که در شکل‌ها و جدول میانگین ویژگی‌ها مشخص است، در ویژگی پولی و تکرار خرید خوشه اول و دوم رفتار تقریباً مشابهی داشته‌اند؛ اما در ویژگی تازگی خرید، خوشه شماره دو، رفتار بهتری نسبت به خوشه شماره یک، به ثبت رسانده است؛ به همین دلیل در رتبه دوم قرار گرفته است.

مرحله چهارم. با توجه به اعضای هر خوشه، سفارش‌های هر خوشه به‌صورت جداگانه بررسی می‌شود؛ بنابراین نیاز است ماتریسی تشکیل شود که تعداد سطرهای آن برابر با تعداد کاربران و تعداد ستون‌های آن برابر با تعداد محصولات باشد. به‌عنوان مثال، برای خوشه صفر ماتریس موردنظر برابر با ۱۶۸۹*۱۰۶۶۵ است و برای مقادیری که ماتریس درصورتی‌که کاربر خرید داشته باشد، مقدار ماتریس برابر با تعداد کالای خریداری‌شده و در غیر این صورت برابر با صفر است.

مرحله پنجم. با استفاده از ضریب همبستگی[1] ارتباط بین کاربر و خوشه محاسبه می‌شود. درصورتی‌که کاربر بخواهد کالایی با کد ۵۲۸۹ را خریداری کند، با توجه به اینکه کاربر در کدام خوشه قرار دارد، پیشنهادهای متفاوتی به او داده می‌شود. برای مثال، کاربر در خوشه صفر خواهان خرید کالایی با کد ۵۲۸۹ است که در این صورت با استفاده از ضریب همبستگی، محاسبه می‌شود با کدام کالاها بیشترین همبستگی را دارد و پیشنهادهای محصول به کاربر نمایش داده می‌شود. به دلیل اینکه تعداد خوشه‌ها برابر سه است، سه جدول همانند جدول ۳، ۴ و ۵، تشکیل می‌شود.

جدول ۳. پیشنهادهای کالا برای خوشه شماره صفر

کد کالا	ضریب همبستگی	قیمت کالا (ریال)
۵۲۸۹	۱	
۲۵۴۳۴۱	۱	۲۶۴۰۰۰۰۰
۸۱۰۶۷	۰/۰۰۰۵۹	۲۲۰۰۰۰
۱۹۹۶۱۶	−۰/۰۰۰۵۹	۳۲۰۰۰۰
۶۸۴۰۵۶	−۰/۰۰۰۵۹	۱۳۵۰۰۰
	میانگین	۲۶۴۰۰۰۰۰

1. Correlation Coefficient

جدول ۴. پیشنهادهای کالا برای خوشه شماره یک

کد کالا	ضریب همبستگی	قیمت کالا (ریال)
۵۲۸۹	۱	
۵۵۰۹۱۶	۰/۷۰	۲۰۰۰۰۰
۵۵۱۵۱۵	۰/۷۰	۱۴۰۰۰۰
۲۷۵۲۶۶	۰/۷۰	۳۱۵۰۰۰۰
۵۴۹۳۵۰	۰/۷۰	۲۸۰۰۰۰
	میانگین قیمت	۷۱۳۷۵۰

جدول ۵. پیشنهادهای کالا برای خوشه‌ی شماره دو

کد کالا	ضریب همبستگی	قیمت کالا(ریال)
۵۲۸۹	۱	
۵۵۰۱۱۸	۰/۷۰	۳۹۰۰۰۰
۷۱۵۹۲۷	۰/۷۰	۳۱۰۰۰۰
۴۹۰۴۶۴	۰/۷۰	۱۱۷۵۰۰۰
۳۳۸۱۴۸	۰/۷۰	۹۸۰۰۰۰
	میانگین قیمت	۹۴۲۵۰۰

در جدول‌های بالا، نمونه کالاهای توصیه‌شده در هر خوشه مشاهده می‌شود که بر اساس ضریب همبستگی از بزرگ به کوچک مرتب شده‌اند و می‌توان کالاهایی که ضریب همبستگی آن‌ها با کالای موردنظر بیش از یک حد آستانه[1] باشد به کاربر به‌عنوان کالاهای پیشنهادی نمایش داد. با در نظر گرفتن حد آستانه برابر ۰/۷ (۷۰ درصد) تنها کالاهایی به کاربر نمایش داده می‌شود که ضریب همبستگی آن کالا بزرگ‌تر مساوی حد آستانه موردنظر باشد. همان‌طور که در جدول‌ها مشخص است، هر کالا با خودش ضریب همبستگی برابر ۱ دارد که این مورد به‌عنوان نمونه کالای پیشنهادی نخواهد بود. در جدول‌های ۳، ۴ و ۵، کُد کالاهایی که حد ضریب همبستگی آن‌ها بیش‌ازحد آستانه است، سبزرنگ و کالاهایی که پایین‌تر از مقدار حد آستانه هستند، قرمز شده‌اند.

مقایسه روش ارائه‌شده با روش‌های قبلی. در روش‌های ارائه‌شده موجود بر روی مجموعه داده «شرکت دیجی‌کالا» برای توصیه محصول ارائه‌ای صورت گرفته نشده است و روش پیشنهادی به‌عنوان رویکرد جدیدی در این مجموعه داده برای توصیه محصول در سایر مجموعه داده‌های مشتریان شرکت‌ها قابل‌اجرا خواهد بود. در روش ارائه‌شده از الگوریتم خوشه‌بندی

1. Threshold

K-means و مدل RFM برای تقسیم استفاده شد؛ همچنین با مطالعه پژوهش‌های موجود، در بیشتر پژوهش‌ها تنها به تقسیم‌بندی مشتریان با استفاده از الگوریتم‌های خوشه‌بندی و مدل RFM پرداخته شده و برای توصیه محصول کاری صورت نپذیرفته است. چند پژوهش به‌عنوان نمونه در ادامه بررسی می‌شود:

در پژوهش (موتمنی و همکاران، ۲۰۱۸) مدلی ترکیبی از الگوریتم‌های تحلیل مؤلفه‌های مستقل، یادگیری منفیلد[1]، خوشه‌بندی و رگرسیون بردار پشتیبان برای پیش‌بینی میزان فروش ارائه شده است [۱۹].

در پژوهش توکلی و همکاران، (۲۰۱۸)، بر روی مجموعه داده شرکت «دیجی‌کالا» تنها برای تقسیم‌بندی مشتریان از روش RFM استفاده شده است [۲۶].

در پژوهش کولکرنی[2] و همکاران، (۲۰۲۲)، تنها یک سیستم توصیه‌گر اجرا شده که از روش ترکیبی برای سیستم توصیه‌گر استفاده شده است و از فیلتر مشارکتی مبتنی بر آیتم و فیلتر مشارکتی مبتنی بر کاربر استفاده شده است. این پژوهش نشان می‌دهد چندین سیستم توصیه‌کننده را می‌توان به‌طور هم‌زمان برای توصیه محصولات به مشتریان پیاده‌سازی کرد [۱۶].

در پژوهش گاسترانسیا[3] و همکاران، (۲۰۲۰)، برای تقسیم‌بندی مشتریان از روش RFM و K-means استفاده شده و برای محاسبه تعداد خوشه‌ی بهینه، ۸ شاخص بررسی شده است [۹].

در پژوهش (شیرول[4] و همکاران، ۲۰۲۱)، از مجموعه داده تجارت الکترونیک بریتانیا که به مدت یک سال از یک خرده‌فروشی آنلاین است استفاده می‌شود و برای تقسیم‌بندی مشتریان از روش RFM و k-means استفاده شده و تعداد خوشه بهینه از روش Elbow محاسبه شده است [۲۵].

در پژوهش چریستی[5] و همکاران، (۲۰۲۱)، از سه الگوریتم خوشه‌بندی K-means و خوشه‌بندی fuzzy C-means و خوشه‌بندی Rm k-means برای تقسیم‌بندی مشتریان بر اساس مدل RFM استفاده شده است؛ سپس کارایی هر یک از الگوریتم‌های خوشه‌بندی با توجه به تعداد تکرارها، متراکم‌بودن خوشه و زمان لازم برای اجرا تحلیل شده است [۱۳].

در پژوهش هارتینی[6] و همکاران، (۲۰۲۰) روش CRISP-DM ارائه شده که در این روش از مدل RFM و خوشه‌بندی k-means استفاده شده است و همانند پژوهش حاضر از روش Elbow برای تعداد خوشه بهینه استفاده شده است. برای استفاده از پروفایل مشتری در مدل از

1. Manifold
2. Kulkarni
2. Gustriansyah
4. Shirole
5. Joy Christy
6. Sari Hartini

طبقه‌بندی‌های Decision Tree و Naive Bayes استفاده شده و مدلی بر اساس RFM+ALC برای تقسیم‌بندی مشتریان به‌کار رفته است [۱۱].

در پژوهش آنجلا[1] و همکاران، (۲۰۲۲) برای تقسیم‌بندی کاربران حمل‌ونقل عمومی برای تحلیل الگوهای مسافران از روش RFM و K-means استفاده شده است. در عبور و مرور، تعداد روزهای پس از آخرین عبور و مرور به‌عنوان پارامتر اخیر، تعداد رفت‌وآمد انجام‌شده در یک دوره معین به‌عنوان پارامتر فرکانس و میزان سود حاصل از تعداد رفت‌وآمد انجام‌شده توسط مسافران در آن دوره به‌عنوان پارامتر پولی استفاده شده است [۱].

۵. نتیجه‌گیری و پیشنهادها

در پژوهش حاضر تلاش شد تا با مطالعه موردی داده‌های فروشگاه آنلاین «دیجی‌کالا» رفتارها و ویژگی‌های مشترک مشتریان بررسی و با توجه به نتایج، یک سیستم توصیه‌گر برای این فروشگاه طراحی شود تا از این طریق علاوه بر درک سلایق و ترجیحات مشتریان که به برنامه‌ریزی بهتر برای خرید محصولات فروشگاه منجر می‌شود، فروش فروشگاه افزایش یابد؛ همچنین باعث افزایش وفاداری مشتریان فروشگاه شود؛ بنابراین این پژوهش با هدف ارائه یک رویکردی تلفیقی از الگوی تراکنشی RFM و الگوریتم خوشه‌بندی K-Means برای تقسیم‌بندی مشتریان و درنهایت توصیه محصول به مشتری صورت گرفت. همان‌طور که ملاحظه شد، در این تحلیل از متغیرهای R، F، M به‌ترتیب به معنی تازگی، تعداد دفعات و مجموع مبالغ مالی در یک دوره‌ی زمانی ۹ روزه برای تقسیم‌بندی استفاده شد. بر اساس اطلاعات جمع‌آوری‌شده برای متغیرهای R,F,M، درنهایت مشتریان به‌وسیله الگوریتم K-Means خوشه‌بندی و به ۳ گروه همگن تقسیم شدند. بر اساس نتایج استخراج‌شده برای هر خوشه می‌توان پیشنهادهایی برای مشتریانی که در یک خوشه مشابه قرار گرفتند، ارائه داد. با توجه به نتایج پژوهش، نتایج حاصل از خوشه‌بندی از نظر ویژگی‌های R,F,M برای خوشه شماره صفر بیشتر است؛ بنابراین با توجه به نتایج مشاهده شد قیمت کالایی که برای خوشه‌ی شماره صفر (مشتریان وفادار) پیشنهاد شده است نسبت به سایر خوشه‌ها بیشتر است که می‌توان برای ترغیب مشتری وفادار به خرید کالاهای با رنج قیمت بالا متناسب با شرایط تخفیف‌های ویژه‌ای در نظر گرفت این سیستم در جهت جذب مشتریان برای جست‌وجو و رسیدن به محصول موردنظر خود بسیار کارساز است.

پیشنهاد می‌شود که در آینده از قوانین انجمنی و همچنین استفاده از شبکه‌های عصبی مانند CNN برای پیاده‌سازی سیستم توصیه محصول استفاده شود و همچنین برای کاهش ابعاد و بهبود نتایج از الگوریتم‌های فراابتکاری استفاده شود. با توجه به پژوهش‌های اخیر، ارائه

1. Angela

متغیرهای اضافی به مدل‌های RFM قدرت پیش‌بینی را افزایش می‌دهد و با توجه به اینکه مدل پایه RFM قادر به تمایز مشتریان بلندمدت و کوتاه‌مدت نیست، تلاش برای بهبود ظرفیت پیش‌بینی مدل برای بهبود نتایج پیشنهاد می‌شود. با توجه به نتایج، استفاده از سیستم‌های توصیه‌گر ترکیبی و استفاده از چند سیستم توصیه‌گر برای پیاده‌سازی سیستم برای بررسی نتایج توصیه می‌شود.

محدودیت‌های پژوهش. در این پژوهش به‌دلیل اینکه مجموعه داده‌های یک شرکت فروش اینترنتی در ایران به‌عنوان مورد کاوی مطالعه شد، محدودیت‌هایی ازجمله محرمانه‌بودن اطلاعات از نظر حفظ حریم خصوصی کاربران وجود داشت؛ از طرف دیگر به‌دلیل محدودیت‌های اعمال‌شده از سمت فروشگاه نیز اطلاعات کامل محصولات در اختیار پژوهشران قرار داده نشد. ازجمله می‌توان به قیمت محصول، ارزش سود هر محصول، میزان تخفیف هر سفارش و برند هر کالا و اطلاعاتی از این قبیل اشاره کرد که در صورت دردست‌داشتن آن‌ها امکان تحلیل‌های بیشتری برای پژوهشگر فراهم می‌شد.

محدودیت‌های زیادی نیز در زمینه‌ی استفاده از داده‌های دنیای واقعی برای تحلیل رفتار مشتریان وجود دارد. به‌دلیل وجود بازار رقابتی، شرکت‌ها به‌سادگی داده‌های مشتریان خود را در اختیار قرار نمی‌دهند و یا در صورت انجام این کار، داده‌های دستکاری‌شده و ناقص در اختیار پژوهشگر قرار می‌دهند. درواقع مدیران شرکت تمایلی برای ارائه اطلاعات به افراد خارج از سازمان ندارند و چون نتایج داده‌کاوی کاملاً به صحت داده‌های اولیه بستگی دارد، این امر ممکن است سبب به‌دست‌آمدن نتایج نامعتبر شود.

عدم همکاری شرکت برای بررسی میزان کارایی سیستم توصیه‌گر طراحی‌شده از دیگر محدودیت‌هایی است که پژوهش حاضر با آن مواجه بود.

تعارض منافع. برای ارائه مطالب و نگارش این مقاله هیچ‌گونه کمک مالی از هیچ فرد، نهاد و سازمانی دریافت نشده است و نتایج و دستاوردهای این مقاله به نفع یا ضرر سازمان یا فردی خاص نخواهد بود. حضور نویسندگان در این پژوهش به عنوان شاهدی بی‌طرف ولی متخصص بوده است و نویسندگان هیچ‌گونه تعارض منافعی ندارند.

منابع

1. Angela, H., Chen, L., Yun-Chia Liang, Wan-Ju Chang, Hsuan-Yuan Siauw, Vanny Minanda. (2022). RFM Model and K-Means Clustering Analysis of Transit Traveller Profiles: A Case Study. *Journal of Advanced Transportation, vol., Article ID 1108105,* 14 pages.
2. Birukov, A., Blanzieri E., & Giorgini P. (2005). Implicit: An agent-based recommendation system for web search[C]. *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems. ACM,* 618-624.
3. Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications, 36(3),* 4176-4184.
4. Choi, S. H., Kang, S., & Jeon, Y. J. (2006). Personalized recommendation system based on product specification values [J]. *Expert Systems with Applications, 31(3),* 607-616.
5. Davidson, I. (2002). *Understanding K-means non-hierarchical clustering.* Computer Science Department of State University of New York (SUNY), Albany.
6. Dibb, S. (1998). Market segmentation: strategies for success. *Marketing Intelligence & Planning, MCB UP Ltd, 16(7),* 394–406.
7. Ernawati, E., Baharin, S. S. K., & Kasmin, F. (2020). A review of data mining methods in RFM-based customer segmentation. *Journal of Physics: Conference Series, (Malang, Indonesia, Volume 1869),* 1-9.
8. Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research, American Marketing Association, 42(4),* 415–430.
9. Gustriansyah, R., Suhandi, N., & Antony, F. (2020). Clustering optimization in RFM analysis Based on k-Means. *Indonesian Journal of Electrical Engineering and Computer Science, 18(1),* 470-475.
10. Han, J., Kamber, M., Pei, J. (2012). *Data Mining: Concept and Techniques* (3rd ed.). USA: Elsevier.
11. Hartini, S., Sigit Kurniawan, W. G., Setiawan, H., Novel, K. (2020). Cosmetics Customer Segmentation and Profile in Indonesia Using Clustering and Classification Algorithm. *Journal of Physics: Conference Series, 1641,* IOP Publishing.
12. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, Elsevier, 31(8),* 651–666.
13. Joy Christy, A., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking – An effective approach to customer segmentation, Journal of King Saud University. *Computer and Information Sciences, 33(10),* 1251-1257.
14. Kahan, R. (1998). Using database marketing techniques to enhance your one-to-one marketing initiatives. *Journal of Consumer Marketing, MCB UP Ltd, 15(5),* 491–493.
15. Khatami FiroozAbadi, M., TaghaviFard, M., Sadjadi, Kh., & Bamdad Soufi, J. (2018). Optimization through simulation to solve the problem of multi-objective allocation of services to the bank's clustered customers. *The Journal of Industrial Management Perspective, 30,* 85-110. (In Persian)
16. Kulkarni, M., & Gulavani, S. (2022). Role of Recommender System for Selling Products in Online Shopping. International Journal of Humanities and Social

Science, “Revival Strategies and Business Policies for Sustainability and Development” on 23rd March 2022, 2582-8568, 660-666.
17. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* 281–297.
18. Michaud, P. (1997). Clustering techniques. Future Generation Computer Systems. *Elsevier, 13(2),* 135–147.
19. Motameni, A., Rezaei, M., & Ehghaghi, M. (2013). Designing a demand prediction model in the ceramic and tile industry. *The Journal of Industrial Management Perspective, 9,* 159-176. (In Persian)
20. Nabizade, F., & Rouhani, S. (2020). Clustering and Prediction Model of Customer Lifetime Value (Case Studies: IRAN National Center for Numbering Goods and Services). *The Journal of Industrial Management Perspective.* (In Persian)
21. Newell, F. (1997). *The New Rules of Marketing: How to Use One-To-One Relationship Marketing to Be the Leader in Your Industry.* McGraw-Hill, Inc., New York.
22. Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review, 13(5-6),* 393-408.
23. Peker, S., Kocyigit, A., & Eren, P. E. (2017). LRFMP model for customer segmentation in the grocery retail industry: a case study. Marketing Intelligence & Planning.
24. Radev, D. R., Fan, W., & Zhang, Z. (2001). Webinessence: A personalized web-based multi-document summarization and recommendation system [J]. *Ann Arbor, 1001,* 48103.
25. Shirole, R., Salokhe, L., & Jadhav, S. (2021). Customer Segmentation using RFM Model and K-Means Clustering. International Journal of Scientific Research in Science and Technology, - 10.32628/IJSRST2183118, 591-597.
26. Tavakoli, M., Molavi, M., Masoumi, V., Mobini, M., Etemad, S., & Rahmani, R. (2018). Customer Segmentation and Strategy Development Based on User Behavior Analysis, RFM Model and Data Mining Techniques: A Case Study. *2018 IEEE 15th International Conference on e-Business Engineering (ICEBE),* 119-126.
27. Thabit, T. H. (2018). The Impact of Customer Relationship Management on Customer Behavior-Case Study of Ooredoo for Telecommunications. *La Revue des Sciences Commerciales, 17(1),* 67-78.
28. Von Reischach, F., Guinard, D., Michahelles, F., et al. (2009). A mobile product recommendation system interacting with tagged products [C]. Pervasive Computing and Communications, PerCom 2009. *IEEE International Conference on. IEEE,* 1-6.
29. Wang, Y.-F., Chuang, Y.-L., Hsu, M.-H., & Keh, H.-C. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications, 26,* 427-434.
30. Wedel, M., Kamakura, & W. A. (2012). Market Segmentation: Conceptual and Methodological Foundations. *Springer Science & Business Media, 8,* 210-215.
31. Wind, Y. (1978). Issues and advances in segmentation research. *Journal of Marketing Research, JSTOR, 15(3),* 317–337.