journal homepage: www.Jnabm.ir

Journal of Intelligent Marketing Management (JIMM)

ISSN: 2783-5405

Presenting a Business Intelligence Model to Improve Customer Relations Performance in Insurance

Seyed Amirhossein Naimaei Mousavi [1], Mohammad Aniseh [*1], Samaneh Monajjem [2]

1- Department of Industrial Management, Faculty of Social Sciences, Imam Khomeini International University, Qazvin, Iran.

2- Ph.D. Student in Future Studies, Faculty of Social Sciences, Imam Khomeini International University, Qazvin, Iran.

Abstract

Objective: This study aims to segment customers of an insurance brokerage unit using business intelligence techniques. Customer segmentation enables managers to gain deeper insights into their clientele and use this information to enhance customer relationship management—a key factor in the insurance industry.
Methodology: The research is applied in nature and follows a descriptive-analytical approach. Data from 32,377 insurance policies belonging to 20,491 individual and organizational customers of the Raees Babaei Insurance Brokerage in Karaj were collected from early 2017 to the end of 2021. After processing and extracting relevant features, the RFM model and customer value matrix were used for segmentation. Clustering techniques including K-Means, Gaussian Mixture Model (GMM), DBSCAN, and hierarchical clustering were applied. The optimal method was selected based on evaluation metrics: Silhouette, Calinski-Harabasz, and Davies-Bouldin indices. The resulting clusters were then analyzed.
Innovation: The study's key innovation lies in integrating business intelligence with multiple clustering methods to identify the most effective customer segmentation strategy, thereby improving organizational interactions with clients.
Findings: The optimal segmentation method was K-Means clustering using normalized data. Customers were grouped into five distinct segments. The analysis also revealed that as the "Recency" variable in the RFM model increases, the proportion of high-value customers in a cluster decreases, while the share of uncertain or low-value customers rises.

Keywords: Business Intelligence, Clustering, Customer Value, RFM

Citation:
Naimaei Mousavi,S. A. Aniseh,M. and Monajjem,S. (2025). Presenting a Business Intelligence Model to Improve Customer Relations Performance in Insurance. Journal of Intelligent Marketing Management, 6(4), 186-211.

Journal of Intelligent Marketing Management (JIMM), 1404, Volume 6, Number 4, Number 30
Publisher: Journal of Intelligent Marketing Management (JIMM)
Type of Article: Scientific Research
https://doi.org/ JABM.3.2.15564.35546.32548
© Authors

Received: 10/06/2025
Received in revised from: 29/07/2025
Accepted: 09/08/2025
Published online: 22/12/2025

journal homepage: www.Jnabm.ir

نشریه مدیریت بازاریابی هوشمند

ISSN: 2783-5405

ارائه یک مدل هوش تجاری به منظور بهبود عملکرد ارتباط با مشتریان بیمه

سید امیرحسین نعیمائی موسوی [۱]، محمد انیسه *[۱]، سمانه منجم [۲]

۱- گروه مدیریت صنعتی، دانشکدهٔ علوم اجتماعی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران.

۲- گروه آینده پژوهی، دانشکدهٔعلوم اجتماعی، دانشگاه بین المللی امام خمینی (ره)، قزوین، ایران.

چکیده

هدف: هدف این پژوهش، بخش‌بندی مشتریان یک واحد کارگزاری بیمه با بهره‌گیری از تکنیک‌های هوش تجاری است. بخش‌بندی مشتریان، مدیران شرکت را قادر می‌سازد تا به بینش عمیق‌تری دررابطه‌با مشتریان خود دست پیدا کنند و از اطلاعات به‌دست‌آمده، در بهبود عملکرد شرکت در حوزه مدیریت ارتباط با مشتریان که یک عنصر کلیدی در صنعت بیمه به شمار می‌آید استفاده کنند.

روش پژوهش: این پژوهش از منظر هدف یک پژوهش کاربردی و از منظر نوع پژوهش، یک پژوهش توصیفی – تحلیلی به شمار می‌آید. به منظور انجام این پژوهش، اطلاعات مرتبط با ۳۲۳۷۷ بیمه‌نامه مربوط به ۲۰۴۹۱ نفر از مشتریان حقیقی و سازمانی واحد کارگزاری بیمه رئیس بابایی کرج در بازه ابتدای سال ۱۳۹۶ تا انتهای سال ۱۴۰۰ گردآوری شدند. پس از پردازش داده‌ها و استخراج مشخصه‌های مرتبط و کاربردی، برای بخش‌بندی مشتریان، از مدل RFM به همراه ماتریس ارزش مشتری استفاده شد. برای تحلیل داده‌ها طبق این مدل‌ها، از روش‌های خوشه‌بندی K-Means ، GMM، DBSCAN و سلسله‌مراتبی استفاده شد. سپس برای انتخاب روش بهینه، نتایج به دست آمده از این روش‌ها با استفاده از روش‌های Silhouette، Calinski-Harabasz و Davies-Bouldin ارزیابی شد. پس از انتخاب روش بهینه، خوشه‌های ایجاد شده مورد تحلیل و بررسی قرار گرفتند. نوآوری اصلی این پژوهش در تلفیق هوش تجاری با روش‌های مختلف خوشه‌بندی و شناسایی بهترین روش بخش‌بندی مشتریان به‌منظور بهبود تعاملات سازمان با مشتریان نهفته است.

یافته‌ها و نتایج پژوهش: نتایج این پژوهش حاکی از این است که روش بهینه برای بخش‌بندی مشتریان، روش خوشه‌بندی K-Means با استفاده از مجموعه‌داده‌های با مقادیر نرمال شده است. مطابق این روش، مشتریان واحد کارگزاری موردمطالعه در ۵ دسته بخش‌بندی می‌شوند. همچنین دریافت شد که با افزایش مقدار متغیر تأخر در مدل تحلیل تأخر، تکرار و مقدار پولی RFM ، به‌تدریج سهم مشتریان ارزشمندتر در یک خوشه کم‌تر شده و سهم مشتریانی که در دسته نامطمئن قرار می‌گیرند بیشتر می‌شود.

کلیدواژه‌ها: هوش تجاری، خوشه‌بندی، ارزش مشتریان، RFM

استناد:

نعیمائی موسوی، سید امیرحسین و انیسه، محمد و منجم، سمانه . (۱۴۰۴). ارائه یک مدل هوش تجاری به منظور بهبود عملکرد ارتباط با مشتریان بیمه. مدیریت بازاریابی هوشمند,۶(۴), ۱۸۶-۲۱۱.

نشریه مدیریت بازاریابی هوشمند،۱۴۰۴، دوره۶، شماره۴، پیاپی ۳۰

ناشر: نشریه مدیریت بازاریابی هوشمند

نوع مقاله: علمی پژوهشی

© نویسندگان

تاریخ دریافت: ۱۴۰۴/۰۳/۲۰

تاریخ ویرایش: ۱۴۰۴/۰۵/۰۷

تاریخ پذیرش: ۱۴۰۴/۰۵/۱۸

تاریخ انتشار: ۱۴۰۴/۱۰/۰۱

https://doi.org/ JABM.3.2.15564.35546.32548

۱- مقدمه

شرکت‌های بیمه‌ای به دلیل فراهم‌کردن خدمات مناسب و بهبود سطح رفاه مردم، به یک جز ضروری در زندگی مردم تبدیل شده‌اند. باتوجه‌به اینکه مشتریان یک عامل کلیدی در کسب درآمد این شرکت‌ها هستند، چگونگی جذب و حفظ مشتریان در شرکت‌های بیمه‌ای از اهمیت بالایی برخوردار است (ارنست و گیجر، ۲۰۲۱). ازاین‌رو، هر گونه تعامل با مشتری از طرف یک شرکت بیمه‌ای حائز اهمیت است. به دلیل وجود رقابت شدید در حوزه بیمه، شرکت‌های فعال در این حوزه به دنبال این هستند تا هم مشتریان جدیدی جذب کنند، هم برای حفظ مشتریان فعلی خود تلاش کنند. علاوه بر این موارد، این شرکت‌ها به دنبال کسب مزیت رقابتی نسبت به رقبای خود نیز هستند؛ بنابراین، پیاده‌سازی صحیح یک سیستم ارتباط با مشتریان به یکی از اولویت‌های اصلی شرکت‌های بیمه‌ای تبدیل شده است. در این راستا، شرکت‌های بیمه‌ای سرمایه‌گذاری عظیمی به‌منظور مدیریت اطلاعات مشتریان خود کرده‌اند؛ ولی به دلیل استفاده از روش‌های سنتی، از حجم زیاد داده‌ها اطلاعات عملیاتی و تصمیم‌یار کمی استخراج می‌شود. (گیکواد، ۲۰۲۰؛ میاه و یه‌او، ۲۰۱۸)؛ بنابراین نیاز به بهره‌جستن از روش‌های با کارایی بالاتر جهت تجزیه‌وتحلیل داده‌ها بیش‌ازپیش احساس می‌شود. متاسفانه در کشور ما علی‌رغم پتانسیل بالای صنعت بیمه در بهبود امنیت خاطر و افزایش رونق اقتصادی، میزان بهره‌گیری از این صنعت بالا نبوده و از تمام قابلیت این صنعت استفاده نشده است (مهدوی و نصیری، ۱۳۹۷). از سوی دیگر، صنعت بیمه در حال حاضر در ایران گرفتار چالش‌های متعددی است که از جمله مهم‌ترین چالش‌ها می‌توان به فناوری قدیمی مورد استفاده جهت جمع‌آوری، دسته‌بندی و تحلیل اطلاعات اشاره نمود. (پژوهشکده بیمه، ۱۴۰۱). متخصصان بیمه برای تجزیه‌وتحلیل نیازهای مشتریان خود به صفحات گسترده پیچیده، تکیه کرده‌اند. این امر سبب شده تا تهیه‌ی گزارش دقیق از داده‌ها با فناوری‌های قدیمی، خسته کننده، زمان‌بر و مستعد خطا گردد (خلیلی دیلمی، آقاسیدمیرزا، ۱۴۰۳).

بنابراین، امروزه شرکت‌های متعددی از جمله شرکت‌های بیمه به دنبال استفاده از روش‌های جدیدتر تجزیه‌وتحلیل داده به‌منظور بهبود عملکرد خود در حوزه‌های مختلف هستند. این شرکت‌ها می‌توانند با استفاده از منابع اطلاعاتی غنی خود از این موضوع بهره ببرند و امکان تصمیم‌گیری بهتری را برای خود فراهم کنند. با بهره‌گیری از تجزیه‌وتحلیل داده‌ها، شرکت‌ها اطلاعات ارزشمندی در مورد مشتریان خود کسب می‌کنند که این اطلاعات موجب دستیابی به مزیت رقابتی نسبت به رقبا می‌شود (ارنست و گیجر، ۲۰۲۱). به کلیت این فرایند، کشف دانش گفته می‌شود. فرایند کشف دانش از طریق داده‌های در دسترس، با کمک سیستم‌های هوش تجاری صورت می‌گیرد.

هوش تجاری فرایند استفاده از فناوری‌های رایانه‌ای به‌منظور گردآوری داده از منابع مختلف، تحلیل داده‌ها و تبدیل تحلیل‌ها به بینش قابل‌استفاده برای مدیران است. مدیران در ادامه از این بینش‌ها باهدف بهبود عملکرد کسب‌وکار و رسیدن به اهداف استفاده می‌کنند. هوش تجاری به شرکت‌ها کمک می‌کند تا به شناخت بهتری از مشتریان خود برسند و رابطهٔ خود را با این افراد بهبود بخشند. بهبود رابطه با مشتریان در نهایت منجر به افزایش فروش و درآمد می‌شود (ماسارو و همکاران، ۲۰۱۹). به عبارتی دیگر، سیستم‌های هوش‌تجاری با ایجاد یکپارچگی سیستم مدیریتی و بهبود فرایندهای کسب‌و کار باعث اثربخشی عملیاتی سازمان‌ها شده و آن‌ها را برای حضور در بازار رقابتی آماده‌تر می‌سازند (محمدی، ۱۴۰۲). از آنجا که همیشه یک شکاف عمیق بین اطلاعات موردنیاز مدیران تجار ی و میزان داده‌هایی که در طول عملیات روزانه شرکت جمع‌آوری می‌شود وجود دارد و به دلیل آن‌که درک نیازهای تمام مشتریان بالقوهٔ یک شرکت و آماده‌سازی برای ارائهٔ خدمات بسته به نیاز منحصربه‌فرد هر مشتری امری بسیار سخت بوده و نیازمند صرف زمان و هزینهٔ بالایی است،

ازاین‌رو لازم است تا شرکت‌ها حوزهٔ فعالیت خود را به چند قسمت تقسیم کنند. با انجام این کار، شناسایی گروه‌های مختلف مشتریان با علایق یا شرایط مشترک امکان‌پذیر می‌شود. خوشه‌بندی مشتریان و تجزیه‌وتحلیل هر کدام از این خوشه‌ها می‌تواند گام مهمی در تحقق اهداف مربوط به ارتباط با مشتریان باشد (پیران و همکاران؛۱۴۰۲؛ جامیوم، ۲۰۲۱). خوشه‌بندی یکی از مهم‌ترین حوزه‌های پژوهش در زمینهٔ داده‌کاوی و هوش تجاری است. این روش برای مطالعهٔ اشخاص بر پایهٔ مجموعه‌ای از مشخصات، باهدف طبقه‌بندی بر اساس مشخصات مذکور است. قاعدهٔ کلی این روش بدین صورت است که اشخاص حاضر در یک دسته‌بندی شباهت بالایی با یکدیگر، و شباهت اندکی با اشخاص حاضر در سایر دسته‌بندی‌ها دارند (ژو، ۲۰۲۰). مطالعه حاضر نیز از روش خوشه‌بندی جهت دسته‌بندی و شناخت الگوهای رفتاری میان مشتریان یک شرکت کارگزاری بیمه به‌منظور ارائهٔ خدمات بهتر بهره برده است. برای تحقق این امر، برای نخستین‌بار محققان این پژوهش ابتدا از روش‌های مختلف خوشه‌بندی شامل K-Means، GMM، DBSCAN و سلسله‌مراتبی استفاده و سپس به کمک روش‌هایی نظیر Silhouette، Calinski-Harabasz و Davies-Bouldin بهینه‌ترین روش خوشه‌بندی را جهت تجزیه‌وتحلیل‌های آتی شناسایی نموده‌اند. باتوجه به مطالب بیان شده، مسئله اصلی این پژوهش ضعف در بهره‌برداری مؤثر از داده‌های مشتریان در شرکت‌های بیمه به‌منظور بهبود ارتباط با مشتری است. با توجه به فقدان بررسی مدل‌های متنوع هوش تجاری مبتنی بر خوشه‌بندی در ادبیات داخلی و خارجی، این تحقیق تلاش دارد با ارائه و ارزیابی یک مدل بهینه از روش‌های خوشه‌بندی، الگویی مؤثر برای تحلیل رفتار مشتریان و ارتقاء سیستم‌های ارتباط با مشتری در صنعت بیمه ارائه داده و به این پرسش اصلی که چگونه یک مدل هوش تجاری می‌تواند در بهبود عملکرد ارتباط با مشتریان شرکت بیمه کمک‌رسان باشد، پاسخ دهد.

۲- مبانی نظری پژوهش

۲-۱- هوش تجاری:

هوش تجاری سیستمی است که با یکپارچه‌سازی ساختار، داده‌ها، ابزارهای تحلیلی، نرم‌افزارها و روش‌های تبدیل داده به اطلاعات مفید، از فرایند تصمیم‌گیری و بهبود موفقیت کسب‌وکار پشتیبانی می‌کند (گیتها و پارام، ۲۰۲۰). هوش تجاری در مدیریت و تجزیه-وتحلیل داده‌های ساختارنیافته و ساختاریافته و مبتنی بر تراکنش مورداستفاده قرار می‌گیرند (سوآدگو و دارمونت، ۲۰۲۱). این مفهوم نقطه تلاقی کسب‌وکار، مدیریت و فناوری اطلاعات محسوب می‌شود. سازمان‌ها با بهره‌گیری از هوش تجاری، اطلاعات تجمیع شده و ساختارمندی را به دست می‌آورند (توربان و همکاران، ۲۰۰۸). دامنه هوش تجاری تا حوزه‌های کاربردی متعددی از جمله مدیریت ارتباط با مشتری، بهینه‌سازی زنجیره تأمین، تجزیه‌وتحلیل مالی، استراتژی‌های بازاریابی، مدیریت ریسک، تحلیل پیش‌بینی و دیگر موارد گسترش می‌یابد (رشیدی، ۱۴۰۲). «در صنعت بیمه هوش تجاری همراه با قابلیت‌های قدرتمند بصری‌سازی داده[1]به بیمه‌گذار کمک می‌کند ضمن اخذ تصمیم‌گیری آگاهانه، از شلوغی کاربرگ‌های پیچیده و سیلوهای داده[2]رهایی یابند» (خلیلی دیلمی، آقاسیدمیرزا، ۱۴۰۳). به‌طورکلی یک سیستم هوش تجاری، بر سه‌پایه انبار داده‌ها، داده‌کاوی و مدیریت عملکرد کسب‌وکار استوار است که به فراخور نیاز سازمان، در هنگام پیاده‌سازی سیستم از این عناصر استفاده می‌شود (موسکوسو و همکاران، ۲۰۱۹).

1 Data visualization
2 Data Silos

۲-۲- خوشه بندی

بخش‌بندی کردن مشتریان به عنوان ابزاری برای متمایز ساختن مشتریان در بازاریابی استراتژیک ضروری است (سان و همکاران، ۲۰۲۱). بخش‌بندی مشتریان این امکان را برای شرکت‌ها فراهم می‌سازد تا به درک بهتری از مشتری‌های خود برسند و بر اساس مشخصه‌های مختلف این مشتریان، استراتژی‌های تفکیک شده‌ای تدوین کنند (وانگ، ۲۰۲۲). جهت بخش‌بندی مشتریان می‌توان از خوشه‌بندی استفاده کرد. در خوشه‌بندی، مجموعه‌ای از اجزای داخل یک دیتاست[1] به بخش‌های مختلفی تقسیم می‌شوند، به‌گونه‌ای که اعضای هر کدام از این بخش‌ها مشخصه‌های مشابهی داشته باشند. بر خلاف طبقه‌بندی، در خوشه‌بندی نتایج مربوط به نتایج هیچ مشخصه‌ای مقداردهی نمی‌شود و با پیاده‌سازی الگوریتم‌های خوشه‌بندی روی دیتاست، مشابهت‌ها بر اساس مشخصه‌های موجود نمایان می‌شوند. از خوشه‌بندی بیشتر در مسائل مربوط به تحلیل سیستم‌های اجتماعی، داروسازی، بازاریابی و باستان‌شناسی استفاده می‌شود. آگاروال و ردی (۲۰۱۴) الگوریتم‌های خوشه‌بندی را در چهار دسته قرار داده‌اند:

خوشه‌بندی مبتنی بر نقطه مرکزی: در این روش، خوشه‌ها بر اساس میزان نزدیکی نقاط با مرکز هر خوشه تشکیل می‌شوند. عضویت هر نقطه در خوشه مربوط به خود با یافتن یک نقطه مرکزی با کم‌ترین فاصله مشخص می‌شود.

خوشه‌بندی مبتنی بر چگالی: مناطق دارای چگالی بالای نمونه تشکیل خوشه می‌دهند. این روش امکان تشکیل خوشه‌هایی با اشکال دلخواه را فراهم می‌کند.

خوشه‌بندی مبتنی بر توزیع: این نوع خوشه‌بندی با فرض تشکیل داده‌ها از چند توزیع (مانند توزیع نرمال یا توزیع گاوسی) انجام می‌شود. با افزایش فاصله میان مرکز توزیع، احتمال تعلق آن نقطه به توزیع مذکور کاهش می‌یابد. از این مدل‌ها زمانی استفاده می‌شود که نوع توزیع داده‌های مشخص باشد.

خوشه‌بندی سلسله‌مراتبی: با درنظرگرفتن هر ردیف داده به‌صورت یک خوشه آغاز شده و سپس در یک فرایند دومرحله‌ای تکرارشونده شامل شناسایی دو خوشه با کمترین فاصله از هم و ادغام دو خوشه با بیشترین شباهت تا جایی که تمام خوشه‌ها در هم ادغام شوند، ادامه می‌یابد خروجی خوشه‌بندی سلسله‌مراتبی به‌صورت یک دندروگرام به نمایش در می‌آید

۲-۳- مدل Recency, Frequency, Monetary (RFM)

مدل تحلیل تأخر، تکرار و مقدار پولی یک روش بخش‌بندی رفتاری قدرتمند و شناخته‌شده در بازاریابی با استفاده از پایگاه‌های داده است. از این مدل معمولاً برای رتبه‌بندی مشتریان بر اساس خریدهای پیشین استفاده می‌شود. مدل RFM در فروشگاه‌های آنلاین، بیمارستان‌ها، شرکت‌های بیمه‌ای و غیره که با مشتریان زیادی سروکار دارد، مورداستفاده قرار می‌گیرد (آگراوال و همکاران، ۲۰۲۰).

۲-۴- فرآیند CRISP-DM

بر اساس تجربه‌های گذشته و نظر محققین و متخصصین این حوزه، فرایندهای متعددی باهدف بیشینه ساختن احتمال موفقیت سیستم‌های هوش تجاری پیشنهاد داده شده‌اند. یکی از رایج‌ترین فرایندهای پیشنهاد شده، فرایندی به نام CRISP-DM است. این فرایند متشکل از شش گام است (توربان و همکاران، ۲۰۱۷).

1 Dataset

۲-۵- روش K-Means :

روش K-Means یکی از پرکاربردترین روش‌های خوشه‌بندی است. مطابق این روش، ابتدا K عدد نقطه به‌عنوان مراکز اولیه خوشه‌ها انتخاب می‌شود. سپس فاصله اقلیدسی محاسبه و هر کدام از داده‌های موجود در دیتاست به نزدیک‌ترین مرکز خوشه تخصیص داده می‌شود. در مرحله بعد، نقاط مرکزی جدید با محاسبه میانگین مقادیر موجود در هر خوشه به‌دست‌آمده و تا زمانی که نقاط مرکزی تثبیت شوند، مراحل تکرار می‌شوند. در انتها با کمک روشی ابتکاری به نام Elbow تعداد بهینه خوشه‌ها برآورد می‌گردد. به این منظور، پس از اجرای خوشه‌بندی، مجموع مربع‌های فواصل با نقطه مرکزی هر خوشه محاسبه و سپس نمودار مرتبط با مجموع مربع‌های فواصل برای بازهٔ اعداد مشخص شده در مرحله اول ترسیم شدند. محل شکست نمودار به‌عنوان تعداد خوشهٔ بهینه برای دیتاست در نظر گرفته می‌شود (یوآن و یانگ، ۲۰۱۹)

۲-۶- روش DBSCAN

روش خوشه‌بندی مبتنی بر چگالی برای مسائل دارای نویز است که بدون نیاز به تعیین پیشینی تعداد خوشه‌ها، قادر به شناسایی خوشه‌های با اشکال نامتعارف و تفکیک نویزها است. این روش با دو پارامتر کلیدی کار می‌کند:

(eps) ε: حداکثر شعاع همسایگی یک نقطه.

MinPts: حداقل تعداد نقاط در همسایگی ε یک نقطه برای تعریف نقطه مرکزی

۲-۷- روش سلسله‌مراتبی

هدف این روش، ایجاد خوشه‌هایی است که اعضای درون آنها مشابه و با اعضای خوشه‌های دیگر متفاوت باشند. اولین گام در خوشه‌بندی به روش سلسله‌مراتبی، محاسبه عدم تشابه هر جفت نقاط در دیتاست است. بدین منظور، ابتدا یک مقیاس فاصله‌ای (مانند فاصله اقلیدسی) برای محاسبه عدم تشابه انتخاب می‌شود. در ادامه دو نقطه‌ای که بیشترین شباهت را دارند تجمیع می‌شوند. این روند تا زمانی که تمامی نقاط دیتاست عضو یک خوشهٔ بزرگ باشند تکرار می‌شوند. در انتهای این فرایند، یک درخت موسوم به دندروگرام[1] ایجاد و ترسیم می‌شود. به‌منظور تعیین میزان نزدیکی دو خوشه با یکدیگر از روش حداقل واریانس Ward که در آن مجموع واریانس بین دو خوشه به حداقل رسیده می‌شود، استفاده شد. در نهایت پس از ترسیم دندروگرام، به‌منظور تعیین تعداد خوشه‌ها، بیشترین فاصله عمودی به‌گونه‌ای که با هیچ خط افقی‌ای برخورد وجود نداشته باشد در نظر گرفته و یک خط افقی از این خط برای کل دندروگرام ترسیم شدند. (جیمز و همکاران، ۲۰۱۳).

۲-۸- روش Guassian Mixture Model (GMM)

یک مدل احتمالاتی است که داده‌ها را حاصل ترکیب خطی چند توزیع گاوسی چندمتغیره با پارامترهای نامعلوم فرض می‌کند. برخلاف K-Means، این روش کوواریانس توزیع‌ها را نیز در نظر می‌گیرد و از الگوریتم EM (حداکثرسازی انتظار) برای تخمین پارامترها استفاده می‌کند. برای محاسبه کوواریانس در روش GMM، چهار گزینه وجود دارد. این گزینه‌ها عبارت‌اند از: الف) کامل: هر توزیع، ماتریس کوواریانس عمومی مربوط به خود را دارد. ب) متصل: تمام توزیع‌ها از یک ماتریس کوواریانس عمومی استفاده می‌کنند ج) قطری: هر توزیع، ماتریس کوواریانس قطری خود را دارد. و) کروی: هر توزیع، ماتریس کوواریانس تکی خود را دارد (پرس،

1 Dendrogram

۲۰۰۷). به منظور یافتن تعداد بهینه‌ی خوشه‌ها در روش GMM، معیار اطلاعات بیزی[1] مورد استفاده قرار می‌گیرد. معیار اطلاعات بیزی به صورت زیر محاسبه می‌شود:

$$BIC = kln(n) - 2ln(L)$$ فرمول (۱)

در این فرمول، BIC بیانگر امتیاز معیار اطلاعات بیزی، k تعداد خوشه‌ها، n تعداد ردیف‌های داده و L تابع احتمال مجموعهٔ موردبررسی است (ویت و همکاران، ۲۰۱۲).

۲-۹- شاخص Silhouette:

مقیاسی برای اندازه‌گیری میزان مشابهت اعضای یک خوشه با یکدیگر، در قیاس با خوشه‌های دیگر است. ضریب Silhouette مقداری بین ۱- و ۱ ارائه می‌کند. هر چه مقدار محاسبه شده به ۱ نزدیک‌تر باشد، بیانگر این است که اعضا تناسب بیشتری با خوشه‌ای که در آن قرار گرفته‌اند دارند. ضریب Silhouette با استفاده از فرمول (۲) محاسبه می‌شود.

$$\boldsymbol{S = \frac{1}{N}\sum_{i=1}^{N}\frac{b_i - a_i}{max(a_i,b_i)}}$$ فرمول (۲)

در این فرمول S بیانگر ضریب Silhouette، a_i بیانگر میانگین عدم تشابه با اعضای دیگر خوشه، b_i بیانگر کم‌ترین میزان عدم تشابه با اعضای سایر خوشه‌ها و N بیانگر تعداد خوشه‌ها است (روساو، ۱۹۸۷) .

۲-۱۰- شاخص Calinski-Harabasz

شاخص Calinski-Harabasz یا شاخص معیار نسبت واریانس، با استفاده از نسبت میان مجموع پراکندگی بین خوشه‌ای برای تمامی خوشه‌ها محاسبه می‌شود. عدد بالای شاخص Calinski-Harabasz بیانگر خوشه‌بندی بهتر است. چرا که اعضای خوشه‌ها به هم نزدیک‌ترند و درعین‌حال، خود خوشه‌ها فاصلهٔ زیادی با هم دارند. طبق این روش پس از محاسبه پراکندگی بیرون و درون خوشه‌ها، شاخص Calinski-Harabasz با کمک فرمول ریر محاسبه می‌شود:

$$\boldsymbol{CH = \frac{\frac{BGSS}{K-1}}{\frac{WGSS}{N-K}} = \frac{BGSS}{WGSS} \times \frac{N-K}{K-1}}$$ فرمول (۳)

در این فرمول $BGSS$ بیانگر پراکندگی بین خوشه‌ای، $WGSS$ بیانگر پراکندگی درون خوشه‌ای و N بیانگر تعداد کل اعضا است (کالینیسکی و هاراباس، ۱۹۷۴).

۲-۱۱- شاخص Davies-Boulein

در این شاخص، کیفیت خوشه‌بندی داده‌ها بر اساس مقادیر و مشخصه‌های موجود در دیتاست موردبررسی محاسبه می‌شود. این روش در انتها یک مقدار مثبت را محاسبه می‌کند. در نهایت هر چه مقدار عدد محاسبه شده کم‌تر باشد، به معنی کیفیت بهتر خوشه‌بندی است (دیویس و بولدین، ۲۰۰۹). ضریب Davies-Bouldin برای خوشه موردبررسی با فرمول زیر بررسی می‌شود:

$$\boldsymbol{\overline{R} = \frac{1}{N}\sum_{i=1}^{N} R_i}$$ فرمول (۴)

۳- پیشینه تجربی پژوهش

به‌کارگیری تکنیک‌های هوش تجاری در مطالعات مختلف مسبوق به سابقه است. تا کنون مطالعات متعددی در این حوزه صورت پذیرفته است. در این راستا، خلیلی‌دیلمی و آقاسیدمیرزا (۱۴۰۳) در مقاله‌ای مروری به‌بررسی نقش هوش‌تجاری در

1 Bayesian information criterion

بازاریابی و عملکرد شرکت‌های بیمه با استفاده از روش کتابخانه‌ای پرداختند. مشفق و همکاران (۱۴۰۳) نیز مدلی جهت ارزیابی عملکرد سازمانی بانک‌ها با کمک ابزارهای هوش‌تجاری BI ارائه نمودند که می‌تواند به تصمیم‌گیران برای اتخاذ تصمیمات استراتژیک یاری رساند. محمد شریفی (۱۳۹۸) با استفاده از روش‌های داده‌کاوی به ارزیابی ارزش مشتریان در حوزه بیمه بدنه اتومبیل پرداخته است. سلیمانی و عاطفت دوست (۱۳۹۸) در تحقیقی مشابه به بررسی هوش تجاری بر تصمیم‌گیری سازمانی با درنظرگرفتن نقش میانجی متغیرهای کیفیت داده و زمینه هوش تجاری پرداختند. امیدی و پورسلیمی (۱۳۹۸) نیز با استفاده از تکنیک AHP فازی، مدلی تعاملی بین هوش بازاریابی و عملکرد صنعت بیمه طراحی نمودند. آن‌ها در این مطالعه با کمک ماتریس مقایسات زوجی، عوامل مؤثر بر ارتباط بین هوش بازاریابی و عملکرد سازمانی را با یکدیگر مقایسه کرده و به آن‌ها وزن داده و در نهایت اولویت هر کدام از عوامل را تعیین نمودند. در مطالعه‌ای دیگر، قاسمی (۱۳۹۷) با ترکیب درخت تصمیم و نرم‌افزار رپیدماینر، چهارچوبی برای شناسایی مشتریان مناسب در بیمه پارسیان ارائه داد. در سطح جهانی نیز مطالعاتی دراین‌رابطه صورت گرفت. راماچانداران و همکاران (۲۰۲۵) در مطالعه‌ای با بهره‌گیری از چارچوب ترکیبی TOE-RBV، چالش‌های پیاده‌سازی سامانه‌های هوش تجاری مبتنی بر هوش مصنوعی در صنعت بیمه مالزی را بررسی کردند. یافته‌های آن‌ها نشان داد که عواملی همچون مقاومت سازمانی، کمبود مهارت‌های تخصصی، زیرساخت‌های فناورانه ناکافی و پیچیدگی‌های مقرراتی از موانع اصلی در مسیر پذیرش این فناوری‌ها هستند. خیاط‌باشی و همکاران (۲۰۲۵) با به‌کارگیری یک مدل زبان بزرگ (LLM) در عملیاتی واقعی از صنعت بیمه، فرایند شناسایی اجزای ادعا را که پیش‌تر به‌صورت دستی انجام می‌شد، اتوماسیون کردند. چوبینگ ژانگ و همکاران (۲۰۲۰) در زمینه ادغام هوش تحلیل کلان‌داده در شرکت‌های چینی، هه و همکاران (۲۰۱۷) مبنی بر مدیریت داده‌های کلان مستخرج از شبکه‌های اجتماعی به‌منظور تصمیم‌گیری برای کسب‌وکار، پومفیووا و بارتکووا (۲۰۱۶) جهت بررسی میزان استفادهٔ شرکت‌های کوچک و متوسط کشور اسلواکی از سیستم‌های اطلاعاتی از قبیل هوش تجاری به‌منظور رسیدن به مزیت رقابتی، اشاره نمودند.

تحقیقات اخیر همچنین بر بهینه‌سازی روش‌های بخش‌بندی مشتریان با مدل‌هایی مانند RFM و خوشه‌بندی تأکید دارند. به‌این منظور، کومار (۲۰۲۵) مقاله‌ای با روش K-Means بر پایهRFM ، مشتریان را به سه گروه کاربردی بخش‌بندی کرده و نتایج قابل اعتماد و تفسیرپذیری بالایی ارائه داده است. لنگ و همکاران (۲۰۲۲) روشی برای وزن‌دهی پویا به مشخصه‌های مدل RFM باهدف بهبود عملکرد بخش‌بندی مشتریان ارائه نمودند. این مقاله نشان داد که تخصیص وزن‌های پویا به مشخصه‌های مدل RFM امری حیاتی بوده و اثر مثبتی بر عملکرد بخش‌بندی مشتریان می‌گذارد. کاور و گابریلچیچ (۲۰۲۲) در پژوهش دیگر از روش‌های بدون ناظر مانند GMM برای بخش‌بندی مشتریان بر اساس الگوهای مصرف استفاده کردند. نتایج پژوهش آن‌ها حاکی از آن بود که زمان مشخص در روز، روز مشخص در هفته، تنوع‌های روزانه و فصلی مواردی هستند که موجب قرارگرفتن مشتریان در خوشه‌های مختلف می‌شوند. ارناواتی و همکاران (۲۰۲۱) نیز به بررسی روش‌های داده‌کاوی مورداستفاده همراه با مدل RFM، به‌منظور ارائهٔ یک چهارچوب بخش‌بندی مشتریان پرداختند. در پژوهش آن‌ها مشخص شد که خوشه‌بندی و به‌تصویرکشیدن داده‌ها، پراستفاده‌ترین روش‌های مورداستفاده از میان هفت روش موردبررسی هستند. در تحقیقی دیگر توسط نانداپالا و همکاران (۲۰۲۰) بخش‌بندی با ترکیب ویژگی‌های جمعیت‌شناختی و تحلیل RFM صورت گرفت که این روش توانست به طور موفقیت‌آمیزی الگوهای شکایت‌ها و میزان شکایت‌ها را بادقت بالایی مشاهده کند.

در ادامه مباحث مطرح شده، تعداد دیگری از پیشینه‌های پژوهش که در ارتباط بامطالعه حاضر می‌باشد، به طور اجمالی در جداول (۱) و (۲) ارائه و بررسی می گردد.

جدول ۱: مطالعات بخش‌بندی مشتریان به کمک هوش تجاری

عنوان	نویسندگان و سال انتشار	الگوریتم‌ها
ارائه یک الگوریتم برای گروه‌بندی انواع محصولات بیمه و کاربران در سیستم توصیه گر بیمه با خوشه بندی مبتنی بر فیلتر مشارکتی ارزیابی عملکرد آن براساس توصیه بیمه	امینی شیرکوهی و یمقانی، ۱۴۰۲	K-means
بخش بندی مشتریان صنعت بیمه با استفاده از الگوریتم خوشه بندی دو مرحله ای مقیاس پذیر (نمونه موردی شرکت بیمه البرز)	پویا و فاضل ترشیزی، ۱۳۹۹	خوشه‌بندی دو مرحله‌ای مقیاس‌پذیر
استفاده از خوشه بندی در مدیریت ارتباط با مشتریان صنعت بیمه	تارخ و اکبری، ۱۳۹۴	K-means
بررسی الگوریتم‌های خوشه بندی برای مشتریان بخش‌های خرده‌فروشی انگستان	John et al, 2023	GMM ،K-means DBSCAN
تحلیل رفتار مشتریان با استفاده از بخش‌بندی فازی شهودگرا	Dogan et al, 2022	K-means FCM
مدل RFM برای تشخیص رفتار خرید مشتریان	Anitha and Patil,2022	K-means

جدول ۲: مطالعات کاربرد هوش تجاری در صنعت بیمه

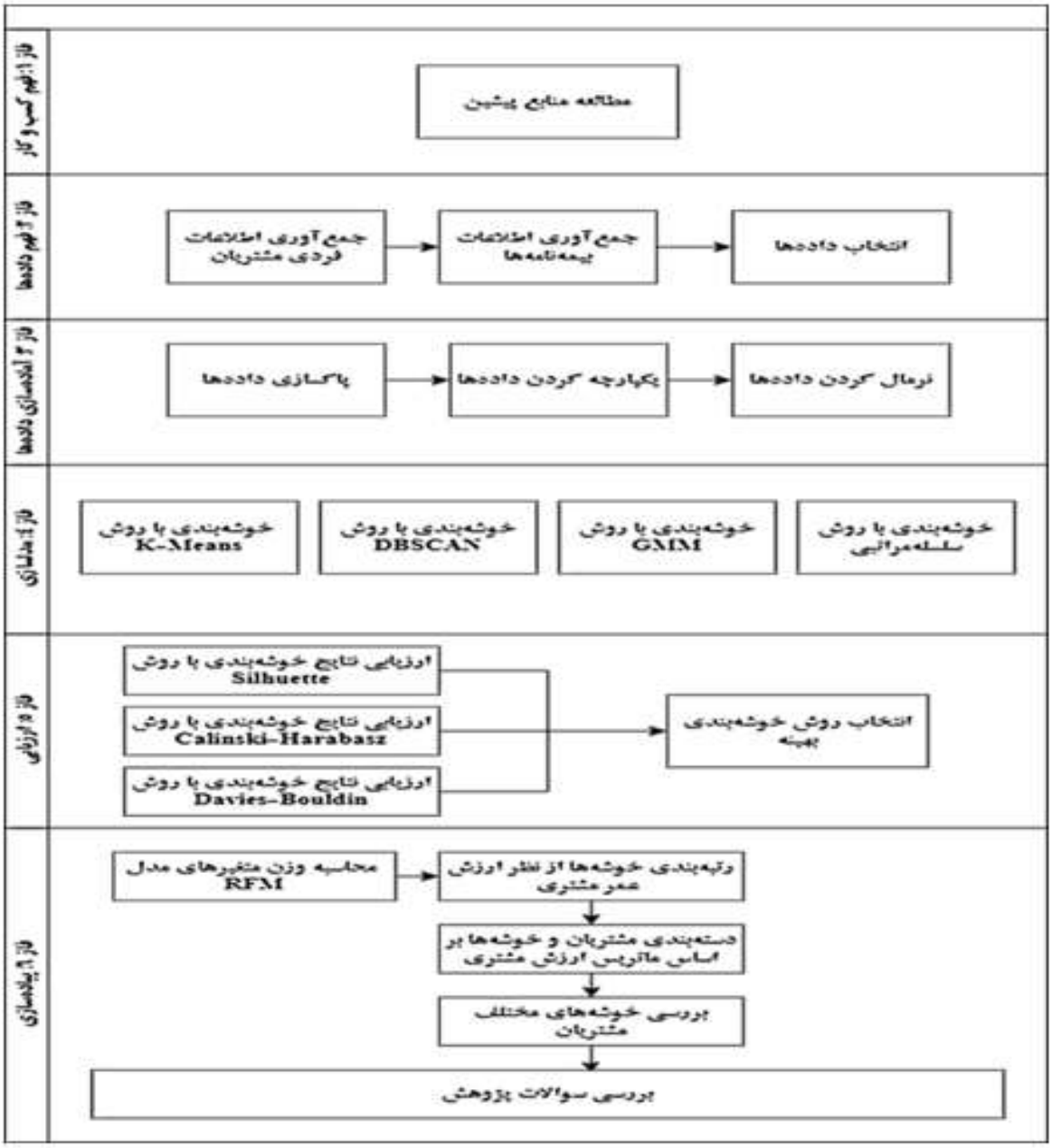
عنوان	نویسندگان و سال انتشار	کاربرد
بهبود تعاملات و تصمیم‌گیری نمایندگان در بیمه با فناوری‌های هوشمند	Das, 2024	بهبود تعاملات و تصمیم‌گیری
تحلیل رفتار مشتری با استفاده از کلان‌داده و یادگیری ماشین	Muradkhanli & Karimov, 2023	شناسایی الگوهای رفتاری مشتریان
رویکرد مبتنی بر یادگیری ماشین برای سپرده خسارت‌ها در بیمه	Baudry, 2019	سپرده خسارت‌ها تجزیه‌وتحلیل خسارت‌ها پیش‌بینی ریسک در بیمه عمر
استراتژی پورتفولیوی پویای بیمه، رویکرد تقویت‌شده یادگیری ماشین	Dehghanpour & Esfahanipour, 2018	پورتفولیوی بیمه‌ای
تحلیل مقایسه‌ای الگوریتم‌های طبقه‌بندی کشف تقلب در بیمه خودرو بر اساس الگوریتم‌های انتخاب ویژگی	Panigrahi and Palkar , 2018	کشف تقلب در بیمه خودرو
مطالعه تکنیک‌های یادگیری ماشین برای پیش‌بینی تقلب در صنعت بیمه	Patil & Godbole, 2018	کشف تقلب

مسئله‌ی قابل توجه در این پژوهش‌ها، تمرکز بیشتر آن‌ها روی مواردی همچون کشف تقلب و پیش‌بینی خسارت بوده و مدیریت ارتباط با مشتریان سهم بسیار کمی را در میان پژوهش‌ها دارد. همچنین مورد مطالعه‌ی پژوهش‌ها بیشتر یک رشته‌ی بیمه‌ای خاص بوده و مطالعه روی تمامی مشتریان و تمرکز بر مواردی همچون واحدهای کارگزاری، در این پژوهش‌ها بسیار کم‌تر دیده می‌شود. همچنین بیشتر این مطالعات به یک یا دو روش خوشه‌بندی اکتفا نمودند. در این پژوهش تلاش شده است تا این مسائل برطرف شود.

۱-۳- استخراج مدل مفهومی پژوهش

پژوهش‌های پیشین نشان می‌دهند که به‌کارگیری روش‌های بخش‌بندی مشتریان از مؤثرترین راه‌های پیاده‌سازی سیستم هوش تجاری در صنعت بیمه است. از میان روش‌های مختلف، الگوریتم‌های خوشه‌بندی و مدل RFM پرکاربردتر هستند. باتوجه‌به تمرکز محدود پژوهش‌های پیشین بر برخی روش‌های خوشه‌بندی، در این پژوهش از الگوریتم‌های مختلف استفاده شده است.

همچنین، به‌جای تمرکز بر یک جنبه از مدلRFM ، چندین جنبه آن بررسی و گروه‌های مشتریان بر اساس ارزش عمر رتبه‌بندی و در ماتریس ارزش مشتری جانمایی شده‌اند. در آخر، گروه‌های مشتریان به‌صورت جداگانه موردبررسی قرار گرفتند تا بینش عمیق‌تری از هر گروه به دست آید. براین‌اساس، مدل مفهومی پژوهش صورت‌گرفته در انتهای این بخش بر اساس مباحث کلی مطرح شده طراحی و در شکل (۱) ارائه شد:



شکل۱: مدل مفهومی پژوهش

۴- روش پژوهش

پژوهش حاضر از نظر هدف، در زمره پژوهش‌های کاربردی قرار می‌گیرد و از منظر روش گردآوری داده‌ها، یک مطالعه توصیفی – تحلیلی محسوب می‌شود. پارادایم حاکم بر تحقیق، پوزیتیویستی است؛ زیرا پژوهش با داده‌های عددی سروکار دارد و با بهره‌گیری از روش‌های آماری، به‌دنبال کشف الگوهای قابل تعمیم در رفتار مشتریان است. استراتژی تحقیق، مطالعه موردی است، چراکه تمرکز بر بررسی عمیق یک واحد مشخص یعنی کارگزاری بیمه رئیس بابایی کرج دارد. پژوهش رویکرد استقرایی را دنبال می‌نماید؛ در این

رویکرد الگوهای رفتاری مشتریان از داده‌های واقعی و بدون فرضیه‌سازی اولیه استخراج شده‌اند. تمام مراحل اجرایی پژوهش با استفاده از زبان برنامه‌نویسی پایتون نسخه ۳ انجام شده است و از کتابخانه‌های متعددی همچون pandas، numpy، matplotlib، seaborn، sklearn، SciPy و Calendar Jalali برای پردازش، مصورسازی و تحلیل داده‌ها استفاده شده است. هدف اصلی، شناسایی الگوهای پنهان در داده‌های مشتریان با استفاده از الگوریتم‌های مناسب، و به‌کارگیری این الگوها برای بهبود عملکرد کارگزاری در حوزه مدیریت ارتباط با مشتریان بوده است.

جامعه آماری این پژوهش شامل ۳۲٬۳۷۷ بیمه‌نامه ثبت‌شده در بازه زمانی فروردین ۱۳۹۶ تا اسفند ۱۴۰۰ متعلق به ۲۰٬۴۹۱ مشتری حقیقی و سازمانی کارگزاری مذکور است .داده‌ها از دو منبع گردآوری شدند: منابع کتابخانه‌ای برای مرور ادبیات نظری و منابع تاریخی شامل اطلاعات واقعی مشتریان برای تحلیل داده‌محور .متغیرهای کلیدی استخراج‌شده عبارت‌اند از: کد ملی یا کد سازمانی، جنسیت، نوع مشتری (حقیقی یا حقوقی)، مبلغ حق بیمه، تاریخ تولد، تاریخ شروع بیمه‌نامه، و نسبت با بیمه‌گذار (در مورد بیمه‌های درمانی).

همان‌گونه که پیش‌تر اشاره شد، روند این پژوهش بر اساس فرایند CRISP-DM که شامل شش گام: ۱) فهم کسب و کار، ۲) فهم داده‌ها، ۳) آماده سازی داده‌ها، ۴) مدل‌سازی ۵) ارزیابی ۶) پیاده سازی بنا شده است. به این منظور پس از مشخص کردن اهداف مطالعه و شناسایی داده‌های مرتبط از منابع موجود، داده‌ها آماده‌سازی می‌شوند. جهت آماده‌سازی داده‌ها شناسه‌های منحصربه فرد با ترکیب کد ملی و کد سازمان ایجاد شده و جنسیت افراد و سازمان‌ها طبقه‌بندی می‌گردد. سپس متغیرهای تاخر زمانی با استفاده از کتابخانه‌های Pandas ، Jalali و Calendar، متغییر تکرار از طریق تراکنش‌های هر مشتری و با کمک کتابخانه Pandas و متغییر پولی نیز مجددا به وسیله کتابخانه و مجموع حق بیمه پرداختی هر مشتری خاص محاسبه می‌شود. در نهایت داده‌ها با فرمول مشخصی نرمال‌سازی و در مراحل بعدی استفاده خواهد شد. داده‌های خام و نرمال شده در جدول (۳) و (۴) نشان داده شده است. در این جدول ستون Recency بیانگر متغیر تاخر، ستون Frequency بیانگر متغیر تکرار و ستون Monetary مربوط به متغیر مقدار پولی است.

جدول ۳: نتیجه آماده‌سازی داده‌های اولیه

	Recency	Frequency	Monetary
0	3	5.000000	1.774548
1	1	4.200000	3.185287
2	2	2.476923	2.825756
3	2	2.415385	2.117917
4	2	2.353846	5.000000
5	2	2.353846	4.048593
6	2	2.169231	2.765380
7	1	2.169231	1.963600
8	1	1.984615	1.603164
9	3	1.984615	1.345826

جدول ۴: مقادیر داده ها پس از نرمال شدن

	Recency	Frequency	Monetary
0	3	66	424096577
1	1	53	1195919500
2	2	25	999218212
3	2	24	611955547
4	2	23	2188758748
5	2	23	1668239036
6	2	20	966186200
7	1	20	527527985
8	1	17	330331262
9	3	17	189540342

در فاز مدل‌سازی، از دیتاست به‌دست‌آمده (خام و نرمال) و چهار روش خوشه‌بندی استفاده خواهد شد. این روش‌ها عبارت‌اند از: ۱) روش K-Means به عنوان یک روش خوشه‌بندی مبتنی بر نقطه مرکزی، ۲) روش خوشه‌بندی سلسله‌مراتبی به عنوان یک روش خوشه‌بندی سلسله‌مراتبی، ۳) روش GMM به عنوان یک روش خوشه‌بندی مبتنی بر توزیع، ۴) روش DBSCAN به عنوان یک روش خوشه‌بندی مبتنی بر چگالی.

۵- پیاده‌سازی

در اولین بخش از آخرین مرحله، وزن هر کدام از مؤلفه‌های مدل RFM محاسبه شد. سپس ۱۴ خبره باسابقه حداقل ۱۰ سال، مقایسه‌های زوجی انجام دادند و ماتریس مقایسه‌ها را تشکیل دادند. پس از محاسبه وزن مؤلفه‌ها، معلومات به‌دست‌آمده در رتبه‌بندی خوشه‌ها بر اساس ارزش عمر مشتریان مورداستفاده قرار گرفت. بدین ترتیب، تمامی خوشه‌های به‌دست‌آمده، از این منظر رتبه‌بندی شدند. فرمول محاسبهٔ ارزش عمر مشتری برای هر کدام از خوشه‌ها به‌صورت زیر است:

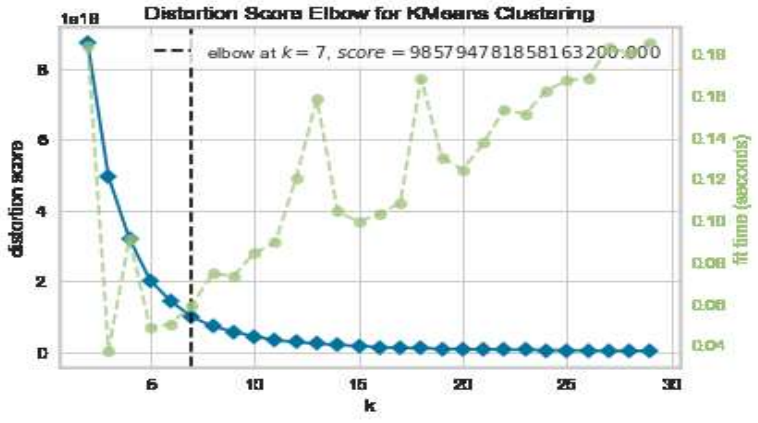
$$CLV_{ci} = NR_{ci} \times WR_{ci} + NF_{ci} \times WF_{ci} + NM_{ci} \times WRM_{ci}$$ فرمول (۵)

در این فرمول CLV_{ci} بیانگر ارزش عمر مشتری خوشهٔ i ام، NR_{ci} بیانگر مقدار میانگین پارامتر تأخر خوشهٔ i ام، WR_{ci} وزن پارامتر تأخر خوشهٔ i ام، NF_{ci} مقدار میانگین پارامتر تکرار خوشهٔ i ام، WF_{ci} وزن پارامتر تکرار خوشهٔ i ام، NM_{ci} مقدار میانگین پارامتر مقدار پولی خوشهٔ i ام و WM_{ci} وزن پارامتر مقدار پولی خوشهٔ i ام هستند (خواجه وند و همکاران، ۲۰۱۱). در دومین بخش فاز پیاده‌سازی، مشتریان و هر کدام از خوشه‌ها بر اساس معیارهای ماتریس ارزش مشتری در چهار دسته دسته‌بندی شدند. در سومین بخش، هر کدام از خوشه‌های ایجاد شده برای مشتریان به صورت جداگانه مورد مطالعه قرار گرفتند.

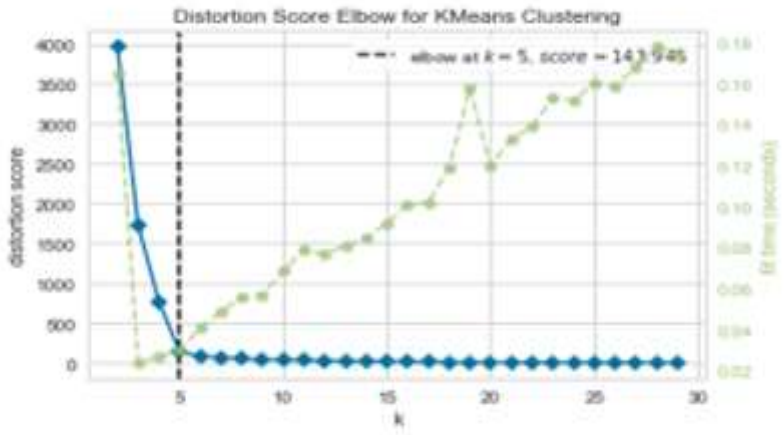
۶- تحلیل یافته‌ها

۶-۱- روش K-Means:

نخستین گام برای اجرای روش K-Means، انتخاب تعداد خوشه‌های بهینه است. با استفاده از روش Elbow، تعداد بهینهٔ خوشه‌ها برای دیتاست موردبررسی برای داده‌های اولیه و نرمال شده به ترتیب ۷ و ۵ خوشه تعیین شد. نتیجه اجرای روش Elbow در داده‌های اولیه و نرمال شده در شکل (۲) و (۳) ارائه شده است.

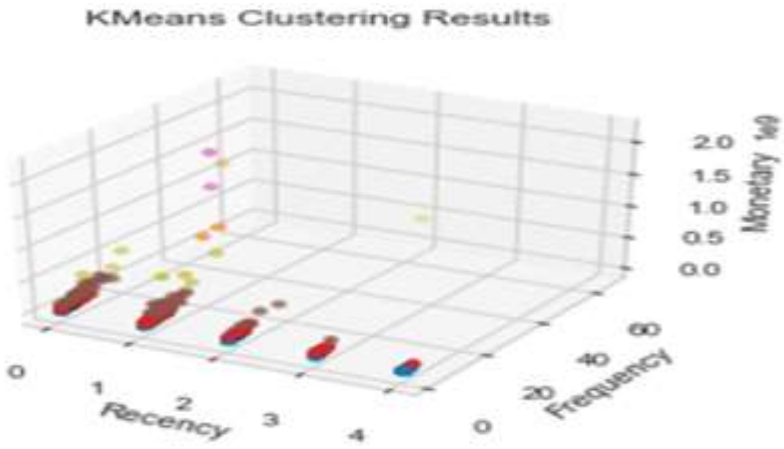


شکل ۲: تعیین تعداد خوشه بهینه برای داده‌های اولیه در روش K-Means

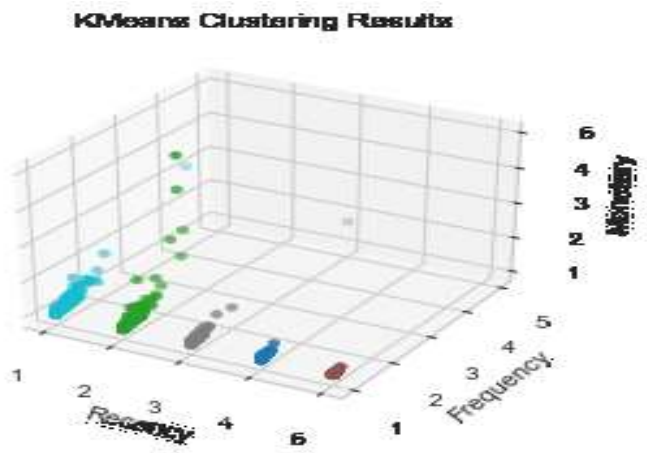


شکل ۳: تعیین تعداد خوشه بهینه برای داده‌های نرمال شده در روش K-Means

پس ازمشخص‌شدن تعداد خوشه‌های بهینه، مراکز جدید خوشه‌ها با درنظرگرفتن نقاط حاضر در هر خوشه مشخص و خوشه‌بندی با روش K-Means اجرا شد. نتایج این خوشه‌بندی در شکل (۴) و (۵) ارائه شده است.



شکل ۴: نتایج مربوط به خوشه‌بندی با روش K-Means برای داده‌های اولیه



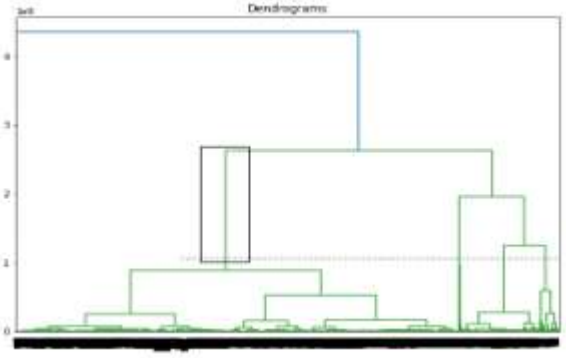
شکل ۵: نتایج مربوط به خوشه‌بندی با روش K-Means برای داده‌های نرمال شده

۶-۲- روش DBSCAN

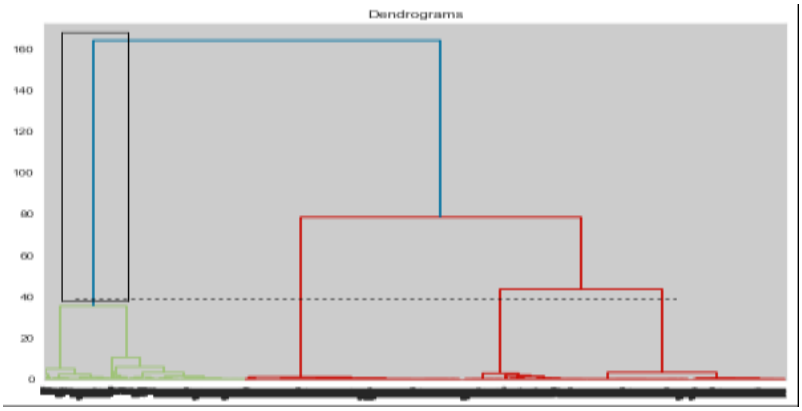
در این روش باتوجه‌به اینکه دیتاست مورداستفاده دارای سه مشخصه است، مقدار این پارامتر برای هر دو داده برابر با ۸ انتخاب شد. پس از پیاده‌سازی با استفاده از مقادیر محاسبه شده، تعداد ۱۵۴ خوشه را برای داده‌های اولیه و ۸۳ خوشه برای داده‌های نرمال شده در نظر گرفته شد. همچنین ۲۲۳۱ و ۱۸۱۵ نقطه به عنوان نقاط نویز برای داده‌های اولیه و نرمال شده شناسایی شدند.

۶-۳- روش سلسله‌مراتبی

دندروگرام این دیتاست با استفاده از فاصله اقلیدسی و روش حداقل واریانس Ward ترسیم گردید. به‌منظور تعیین تعداد خوشه‌ها، بزرگ‌ترین فاصله عمودی بدون برخورد با خطوط افقی به‌عنوان مبنا در نظر گرفته شده که این خط در شکل‌ها با یک حاشیه سیاه مشخص شده است. سپس خطی افقی از این خط ترسیم گردیده و تعداد نقاط برخورد این خط با خطوط عمودی، بیانگر تعداد خوشه‌ها در این روش خواهد بود. مطابق شکل (۶) و (۷) برای داده‌های اولیه و نرمال شده، تعداد ۴ خوشه مطابق تعداد نقاط برخورد در نظر گرفته شد.



شکل ۶: تعیین تعداد مناسب خوشه‌ها برای داده‌های اولیه در روش سلسله‌مراتبی



شکل ۷: تعیین تعداد مناسب خوشه‌ها برای داده‌های نرمال شده در روش سلسله‌مراتبی

۶-۴- روش GMM

باتوجه‌به سازوکار معیار اطلاعات بیزی جدول‌های شمار ۵ و ۶، هر چه مقدار این معیار عدد پایین‌تری را نمایش دهد، تعداد اجزا و کوواریانس مرتبط با آن مقدار، موارد مناسب‌تری برای خوشه‌بندی داده‌های مورد بررسی به شمار می‌آیند. لذا با استناد به مقادیر جداول، تعداد بهینه‌ی خوشه‌ها در روش GMM برای داده‌های اولیه و نرمال شده به ترتیب عدد ۶ و ۱۵ بوده و روش بهینه‌ی محاسبه‌ی کوواریانس، روش کامل تعیین شده است.

جدول ۵: نتایج مربوط به امتیاز معیار اطلاعات بیزی برای داده‌های اولیه

کوواریانس کامل		کوواریانس متصل		کوواریانس قطری		کوواریانس کروی	
امتیاز BIC	اجزا	امتیاز BIC	اجزا	امتیاز BIC	اجزا	امتیاز BIC	اجزا
457786.7	1	457786.7	1	467011.2	1	1143369	1
425042.8	2	448972.9	2	432142.8	2	1083447	2
342860.5	3	444929.2	3	343637.1	3	1051844	3
340431.7	4	444201	4	340815.1	4	1047457	4
340381.9	5	439564.3	5	311929.9	5	1045563	5
308202.5	6	435673.7	6	311791.3	6	950834.6	6

جدول ۶. نتایج مربوط به امتیاز معیار اطلاعات بیزی برای داده‌های نرمال شده

کوواریانس کامل		کوواریانس متصل		کوواریانس قطری		کوواریانس کروی	
امتیاز BIC	اجزا	امتیاز BIC	اجزا	امتیاز BIC	اجزا	امتیاز BIC	اجزا
-16139.1	1	-16139.1	1	-6914.56	1	71638.9	1
-48978	2	-19177.4	2	-41750.4	2	35609.12	2
-61005.5	3	-22724.6	3	-53409.2	3	22881.3	3
-133842	4	-21702.4	4	-126831	4	-75962.7	4
-174365	5	-143774	5	-166842	5	-93812.4	5
-182820	6	-152233	6	-175658	6	-112199	6
-183755	7	-153722	7	-176921	7	-114107	7
-182990	8	-157399	8	-185106	8	-124705	8
-190249	9	-159876	9	-186390	9	-126582	9
-190733	10	-160358	10	-187158	10	-128313	10
-191312	11	-160693	11	-188456	11	-133234	11
-205838	12	-160656	12	-207152	12	-153832	12
-205922	13	-162132	13	-207546	13	-154724	13
-205858	14	-164638	14	-207153	14	-151616	14
-213529	15	-169026	15	-207470	15	-154742	15
-205775	16	-169478	16	-205436	16	-154851	16
-205762	17	-170162	17	-205431	17	-154848	17
-205685	18	-170778	18	-205403	18	-154868	18
-206148	19	-171809	19	-206248	19	-154134	19
-206089	20	-172417	20	-206205	20	-154294	20

۵۶- نتایج خوشه‌بندی

روش‌های مختلف ارزیابی کیفیت خوشه‌بندی، شاخص‌های متفاوتی برای این کار دارند. در روش Silhouette هر چه ضریب Silhouette به عدد ۱ نزدیک‌تر باشد، کیفیت خوشه‌بندی بالاتر ارزیابی می‌شود. با استفاده از این معیار، روش K-Means در داده‌های نرمال شده بالاترین کیفیت خوشه‌بندی را ارائه می‌کند.

در روش Calinski-Harabasz بالاتر بودن مقدار شاخص Calinski-Harabasz بیانگر کیفیت بهتر خوشه‌بندی است. در این شاخص ارزیابی نیز روش خوشه‌بندی K-Means با استفاده از داده‌های نرمال شده کیفیت بسیار بیشتری را در قیاس با سایر روش‌های خوشه‌بندی ارائه می‌کند.

در روش Davies-Bouldin تنها مقادیر مثبت ایجاد می‌شوند. هر چه شاخص Davies-Bouldin به عدد صفر نزدیک‌تر باشد، روش به‌کاررفته دارای دقت و کیفیت بالاتری است. در این شاخص، روش K-Means با استفاده از داده‌های نرمال شده نزدیک‌ترین مقدار به صفر را داشته و به‌عنوان دقیق‌ترین روش شناسایی شده است.

باتوجه‌به نکات ارائه شده در این بخش، روش خوشه‌بندی K-Means با استفاده از داده‌های نرمال شده در هر ۳ شاخص ارزیابی مورداستفاده، به‌عنوان بهترین روش شناسایی شده است. ازاین‌رو، از این روش برای پیاده‌سازی تجزیه‌وتحلیل‌های نهایی روی داده‌ها و ارائهٔ نتایج استفاده می‌شود. نتایج کلی این روش‌های ارزیابی در جدول ۷ ارائه شده است.

جدول۷: نتایج کلی ارزیابی روش‌های خوشه‌بندی

	Silhouette		Calinski-Harabasz		Davies-Bouldin	
	نرمال شده	اولیه	نرمال شده	اولیه	نرمال شده	اولیه
K-Means	۰,۹۳	۰,۶۸	۳۲۵۲۳۲	۳۰۱۰۳	۰,۱۶	۰,۴۴
DBSCAN	۰,۷۰	۰,۵۲	۱۴۹۳۱	۱۳,۵۵	۱,۳۸	۱,۴۰
سلسله‌مراتبی	۰,۸۷	۰,۷۱	۷۷۴۴۸	۱۵۳۲۷	۰,۹۶	۱,۲۳
GMM	۰,۷۵	۰,۲۹	۱۴۴۱۰۵	۶۳۳۸	۰,۲۹	۰,۵۲

۶-۶- محاسبه وزن مولفه‌های مدل RFM

پس از جمع‌آوری داده‌های تکمیل شده توسط افراد خبره به‌صورت ماتریس مقایسه‌های زوجی، از نرم‌افزار Exper Choice به‌منظور بررسی سازگاری مقایسه‌ها و محاسبه وزن نهایی مشخصه‌ها استفاده شد. درجه ناسازگاری مقایسه‌های زوجی ۰,۰۰۰۰۹ به دست آمد که کم‌تر از ۰,۱ بوده و قابل‌قبول به شمار می‌آید. در نهایت، وزن مشخصه‌های R، F و M به ترتیب ۰,۰۶۱، ۰,۴۷۲ و ۰,۴۶۷ به دست آمد. نتایج این پژوهش نشان می‌دهند که مشخصه تکرار خرید یا تمدید بیمه‌نامه دارای بیشترین اهمیت بوده و پس از آن مشخصه حجم پولی خرید با اختلاف ۰,۰۰۵ دارای بالاترین اهمیت است. در نهایت نیز مدت‌زمان سپری شده از آخرین خرید یا تمدید بیمه‌نامه کم‌ترین اهمیت را در میان مشخصه‌های مدل RFM دارد.

۶-۷- رتبه‌بندی خوشه‌ها از نظر ارزش عمر مشتری

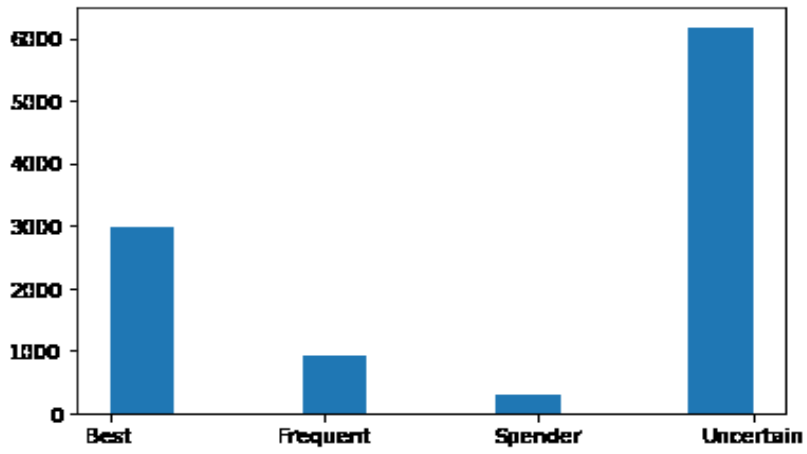
با محاسبهٔ ارزش عمر مشتری برای هر کدام از خوشه‌ها، می‌توان خوشه‌ها را از منظر ارزش عمر مشتری رتبه‌بندی کرد. نتایج مربوط به رتبه‌بندی خوشه‌ها از منظر ارزش عمر مشتری در جدول ۸ ارائه شده است.

جدول۸: نتایج رتبه‌بندی خوشه‌ها از منظر ارزش عمر مشتری

خوشه	ارزش عمر مشتری	درصد جمعیت	رتبه
۱	۱,۱۷۱	۱۳,۲	۳
۲	۱,۰۱۴	۲۹,۳	۵
۳	۱,۴۰۰	۹,۲	۱
۴	۱,۲۵۴	۱۸,۱	۲
۵	۱,۰۹۲	۳۰,۲	۴

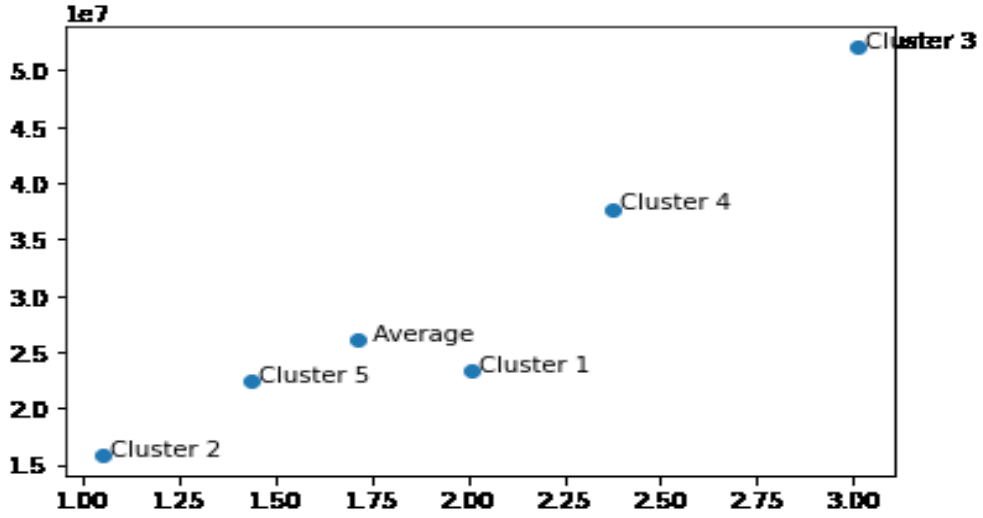
۶-۸- دسته‌بندی اعضای خوشه‌ها بر اساس ماتریس ارزش مشتری

مطابق با مؤلفه‌های ماتریس ارزش مشتری، تمامی مشتریان در ۴ دسته گروه‌بندی شدند. به‌منظور دستیابی به مقادیر صحیح، ابتدا میانگین دو پارامتر تکرار و مقدار پولی محاسبه شدند. میانگین پارامتر تکرار برابر با ۱,۷۱ و میانگین پارامتر مقدار پولی برابر با ۲۶۱۰۹۷۶۲,۳۲ محاسبه شدند. با درنظرگرفتن این مقادیر، مشتریان در ۴ دستهٔ بهترین (هر دو پارامتر بالاتر از میانگین)، سخاوتمند (پارامتر مقدار پولی بالاتر از میانگین و پارامتر تکرار پایین‌تر از میانگین)، پرتکرار (پارامتر مقدار پولی پایین‌تر از میانگین و پارامتر تکرار بالاتر از میانگین) و نامطمئن (هر دو پارامتر پایین‌تر از میانگین) قرار داده شدند. نتایج این دسته‌بندی در شکل (۸) ارائه شده است. ۶۱۷۳ مشتری در دسته‌ی نامطمئن، ۲۹۸۶ مشتری در دسته‌ی بهترین، ۹۰۷ مشتری در دسته‌ی پرتکرار و ۲۷۸ مشتری در دسته‌ی سخاوتمند قرار گرفتند.



شکل۸: نتایج دسته‌بندی مشتریان بر اساس ماتریس ارزش مشتری

علاوه بر اعضای خوشه‌ها، خود خوشه‌ها نیز بر اساس ماتریس ارزش مشتری دسته‌بندی شدند. بدین منظور، میانگین پارامترهای تکرار و مقدار پولی به‌صورت جداگانه برای هر کدام از خوشه‌ها محاسبه و با میانگین کل داده‌ها مقایسه شد. نتایج این دسته‌بندی در شکل (۹) ارائه شده است. مطابق با نتایج این دسته‌بندی، خوشه‌ی اول در دسته‌ی پرتکرار، خوشه‌ی دوم در دسته‌ی نامطمئن، خوشه‌ی سوم در دسته‌ی بهترین، خوشه‌ی چهارم در دسته‌ی بهترین و خوشه‌ی پنجم در دسته‌ی نامطمئن قرار می‌گیرند. نکته‌ی حائز اهمیت در این دسته‌بندی، عدم قرارگیری هیچ یک از خوشه‌ها در دسته‌ی سخاوتمند است.



شکل۹: نتایج دسته‌بندی خوشه‌ها در ماتریس ارزش مشتری

۶-۹- بررسی خوشه‌ها

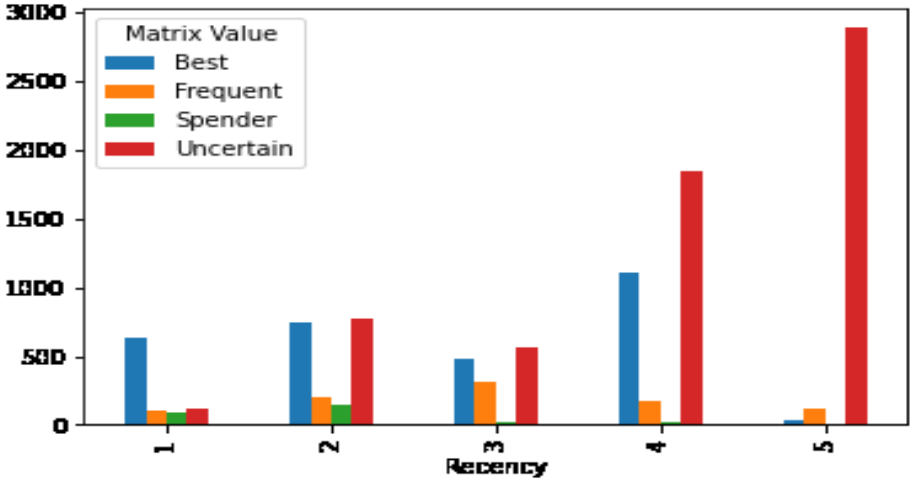
هر کدام از پنج دسته مشتریان ایجاد شده، از دو منظر سهم جنسیتی در هر کدام از خوشه‌ها و وضعیت مشتریان (بهترین، سخاوتمند، پرتکرار، نامطمئن) از منظر ماتریس ارزش مشتری بررسی شدند. نتایج مجتمع شدهٔ این بررسی‌ها در جدول (۹) ارائه شده است. مطابق نتایج به‌عمل‌آمده خوشه سوم با ۳٬۰۲۹ مشتری (۹۰٪ مرد) به‌عنوان ارزشمندترین خوشه از نظر ارزش عمر مشتری شناخته شد و در ماتریس ارزش مشتری در دسته «برتر» قرار گرفت. خوشه چهارم نیز با رتبه دوم ارزش عمر و جایگاه «برتر» در ماتریس، عملکرد مشابهی داشت. در مقابل، خوشه دوم (کم‌ارزش‌ترین) و خوشه پنجم (رتبه چهارم) در دسته «نامطمئن» ماتریس جایگزینی شدند. خوشه اول با داشتن ۱۸۳۸ مشتری که عمدتاً در دسته بهترین و نامطمئن قرار دارند سومین خوشهٔ ارزشمند داده‌های موردبررسی، از منظر ارزش عمر مشتری است.

توزیع جنسیتی در بین خوشه‌های مختلف متفاوت است. در تمام خوشه‌ها سهم مردان غالب بود (بین ۷۴٪ تا ۹۰٪) و به‌ویژه خوشه‌های ارزشمندتر تعداد مشتریان مرد به لحاظ قابل‌توجهی بالاتر است. همچنین درصد کمی از مشتریان به‌صورت سازمانی تعریف شده‌اند.

جدول ۹. نتایج مرتبط با مشخصات مشتریان در هر کدام از دسته‌بندی‌ها

خوشه	جنسیت	نوع مشتری			
		بهترین	سخاوتمند	پرتکرار	نامطمئن
۱	مرد	۹۷۴	۱۳	۱۳۷	۱۴۳۶
	زن	۱۲۵	۵	۳۳	۴۰۱
	سازمانی	۳	۰	۱	۱
۲	**مرد**	۵۱۲	۵۸	۷۷	۸۷
	زن	۱۲۴	۳۳	۲۸	۳۱
	سازمانی	۲	۱	۰	۰
۳	مرد	۲۷	۰	۹۳	۲۵۸۹
	زن	۴	۱	۱۸	۲۹۱
	سازمانی	۰	۰	۳	۳
۴	مرد	۳۵۶	۱۳	۲۴۲	۴۳۳
	زن	۱۱۰	۳	۶۸	۱۳۱
	سازمانی	۹	۰	۰	۰
۵	مرد	۵۷۴	۱۰۱	۱۶۲	۶۳۶
	زن	۱۶۰	۵۰	۴۵	۱۳۴
	سازمانی	۶	۰	۰	۰

همچنین با افزایش مقدار متغیر تاخر در مدل RFM، به‌تدریج از سهم مشتریان قرار گرفته در بخش‌برتر کاسته شده و به مشتریان نامطمئن افزوده می‌شود. نتایج مربوط به این روند در شکل (۱۰) ارائه شده است.



شکل ۱۰: ارتباط میان مقدار متغیر تاخر با سهم مشتریان برتر و نامطمئن

۷- نتیجه‌گیری

بیمه صنعتی با کاربرد بسیار گسترده است که با رشد خود موجب ایجاد امنیت خاطر و رونق اقتصادی می‌شود. بااین‌وجود، میزان بهره‌گیری از این صنعت در کشور ما بالا نبوده و از تمام پتانسیل این صنعت در کشور استفاده نشده است. به‌طوری‌که در سال ۲۰۱۷، تقاضای سرانه بیمه در ایران ۱۷ درصد مقدار مشابه در کل جهان بوده است. ازاین‌رو هر گونه تعامل با مشتری از طرف شرکت‌های بیمه به‌منظور بهره‌بردن از فرصت‌های مالی و اقتصادی این صنعت، حائز اهمیت است. پژوهش حاضر باهدف بهبود عملکرد ارتباط با مشتریان یک شرکت کارگزاری بیمه با ارائه یک مدل هوش تجاری صورت پذیرفته است. تمامی مراحل پژوهش با استفاده از زبان برنامه‌نویسی پایتون نسخه ۳ انجام شده است. جهت پیشبرد مراحل این پژوهش، از کتابخانه‌های مختلفی همچون کتابخانه pandas، numpy، matplotlib، seaborn، sklearn، SciPy، Jalali Calendar استفاده شده است. جهت گردآوری داده‌ها، اطلاعات مربوط به مشتریان کارگزاری بیمه رئیس بابایی واقع در شهر کرج، در بازهٔ فروردین سال ۱۳۹۶ تا اسفند سال ۱۴۰۰ جمع‌آوری شدند. از این داده‌ها برای بخش‌بندی مشتریان با استفاده از مدل RFM استفاده شد. این داده‌ها پس از پاک‌سازی و انتخاب داده‌های موردنیاز، در دو نسخهٔ اولیه و نرمال شده به مرحلهٔ تجزیه‌وتحلیل داده‌ها منتقل شدند. در مرحلهٔ تجزیه‌وتحلیل داده‌ها، این دو مجموعهٔ داده با استفاده از ۴ روش خوشه‌بندی K-Means، DBSCAN، GMM و روش سلسله‌مراتبی موردبررسی قرار گرفتند. به‌منظور ارزیابی کیفیت خوشه‌بندی در روش‌های مورداستفاده، از ۳ شاخص Silhouette، Calinski-Harabasz و Davies-Bouldin استفاده شد. نتایج این ارزیابی‌ها حاکی از آن است که مستقل از روش خوشه‌بندی به‌کاررفته، داده‌های نرمال شده نسبت به داده‌های اولیه کیفیت خوشه‌بندی بهتری از خود نشان می‌دهند. در ادامه دریافت شد که روش DBSCAN برای خوشه‌بندی این جنس از داده‌ها روش مناسبی به شمار نمی‌آید. چرا که این روش در شناسایی داده‌های موسوم به Outlier عملکرد ضعیفی داشته و این داده‌ها را به‌عنوان نویز در نظر می‌گیرد. در نهایت روش K-Means با استفاده از داده‌های نرمال شده به‌عنوان روش خوشه‌بندی بهینه شناسایی شد و از آن در قسمت استخراج نتایج استفاده شد.

پس از انتخاب روش بهینه، پنج خوشه شناسایی شد و امتیاز هر کدام از خوشه‌های ایجاد شده با استفاده از وزن هر کدام از پارامترهای مدل RFM محاسبه و خوشه‌ها رتبه‌بندی شدند. هر کدام از پنج دسته مشتریان ایجاد شده، از دو منظر موردمطالعه قرار گرفتند، ابتدا تعیین

شد که سهم جنسیتی در خوشه‌ها به چه صورت است و سپس وضعیت مشتریان از منظر ماتریس ارزش مشتری مشخص شد. بنا بر نتایج به‌عمل‌آمده، خوشه سوم و چهارم به ترتیب ارزشمندترین خوشه شناخته شدند. هر دو این خوشه‌ها بر نقش پررنگ مشتریان مرد در ایجاد ارزش اقتصادی برای شرکت تأکید دارند، هرچند حضور چشمگیر مردان در بخش «نامطمئن» خوشه سوم (۲،۵۸۹ نفر) نشان‌دهنده نیاز به مدیریت ریسک این گروه است. در مقابل، خوشه دوم (کم‌ارزش‌ترین) و خوشه پنجم (رتبه چهارم) به دلیل تأخر بالای خرید یا وفاداری ضعیف، در دسته «نامطمئن» طبقه‌بندی شده‌اند که لزوم اجرای استراتژی‌های بازگرداندن مشتریان را پررنگ می‌کند .خوشه اول نیز با ترکیب متناقض مشتریان «بهترین» و «نامطمئن»، نیازمند تحلیل عمیق‌تر یا تقسیم به زیر خوشه‌ها برای شناسایی الگوهای رفتاری خاص است. از سوی دیگر، توزیع جنسیتی در تمام خوشه‌ها نشان‌دهنده غلبه مردان (۷۴٪ تا ۹۰٪) است که احتمالاً بازتابی از تناسب محصولات و خدمات با نیازهای مردان یا سوگیری در استراتژی‌های بازاریابی است. این نابرابری جنسیتی، فرصتی برای تنوع‌بخشی به پایگاه مشتریان از طریق جذب هدفمند زنان و بهبود کمپین‌هاست. همچنین، سهم ناچیز مشتریان سازمانی (کمتر از ۱٪) حاکی از تمرکز فعلی شرکت بر مشتریان فردی است که توسعه‌بخش B2B و ارائه مزایای ویژه به سازمان‌ها می‌تواند به رشد پایدار منجر شود.

ازاین‌رو، می‌توان گفت که تمرکز بر مشتریان مرد باارزش اقتصادی بالا، هسته اصلی سودآوری شرکت است، اما تنوع‌بخشی به پایگاه مشتریان (به‌ویژه جذب زنان و سازمان‌ها)، بازتعریف استراتژی‌های خوشه‌بندی برای کاهش تناقضات، و کاهش تأخر خرید از طریق مدل RFM، مسیرهای کلیدی برای بهبود عملکرد بلندمدت و کاهش ریسک‌های مالی به شمار می‌روند.

می‌توان با کمک اطلاعات حاصل از این تحقیق و به‌وسیله‌ی راهبردهایی نظیر برقراری تماس با مشتریان دسته‌ی نامطمئن و تلاش برای تعامل با این مشتریان و ترغیب به خریداری خدمات بیشتر با استفاده از بسته‌های تخفیفی مختلف، مشخص کردن مشتریانی که مدت زمان زیادی از آخرین خریدشان می‌گذرد و قطع منابع اختصاص داده شده به این مشتریان و همچنین تدوین برنامه‌های بیشتر برای سودده‌ترین مشتریان به منظور حفظ این افراد، به بهبود عملکرد تعاملی با مشتری و افزایش بازده اقتصادی دست یافت.

وجود پیاده‌سازی دقیق و موفق الگوریتم‌های خوشه‌بندی متنوع در این پژوهش، یکی از محدودیت‌های مهم، عدم استفاده از روش‌های ارزیابی بیرونی جهت بررسی پایداری خوشه‌ها است. در این پژوهش از معیارهای درونی نظیرSilhouette ، Davies–Bouldinو Calinski–Harabaszبرای ارزیابی کیفیت خوشه‌بندی استفاده شده است؛ با این حال، برای افزایش اعتبار و تعمیم‌پذیری نتایج، پیشنهاد می‌شود در پژوهش‌های آتی:

از روش Cross-validation در خوشه‌بندی) نظیر (K-fold clustering برای ارزیابی پایداری ساختار خوشه‌ها در بخش‌های مختلف داده استفاده شود همچنین. بخشی از داده‌ها به‌عنوان داده‌های خارج از نمونه (Out-of-sample) در نظر گرفته شده و برای ارزیابی کیفیت خوشه‌بندی روی داده‌های جدید و پیش‌بینی‌نشده به کار گرفته شود.

از شاخص‌های ارزیابی بیرونی مانند Adjusted Rand Index یا Normalized Mutual Information برای مقایسه بین نتایج حاصل از اجراهای مختلف الگوریتم‌ها بهره گرفته شود.

جمع‌آوری داده از چندین واحد کارگزاری بیمه در سراسر کشور، مقایسه‌ی نتایج به دست آمده مربوط به هر کدام از واحدها و بررسی تاثیر موقعیت جغرافیایی روی نتایج به‌کارگیری این روش‌ها می‌تواند دقت، پایداری و تعمیم‌پذیری الگوی بخش‌بندی مشتریان را به‌شکل معناداری ارتقا دهد.

منابع:

امیدی، انوش و پورسلیمی، مجتبی. (۱۳۹۸). طراحی مدل تعاملی هوش بازاریابی و اثربخشی عملکرد صنعت بیمه با بهره‌گیری از تکنیک AHP فازی مطالعهٔ موردی: بیمه پاسارگاد. *تحقیقات بازاریابی نوین*. ۹(۲)، ۱۴۹-۱۶۸.

امینی شیرکوهی,مرضیه؛ یمقانی,محمدرضا. (۱۴۰۲). ارائه یک الگوریتم برای گروه بندی انواع محصولات بیمه و کاربران در سیستم توصیه گر بیمه با خوشه بندی مبتنی بر فیلتر مشارکتی ارزیابی عملکرد آن براساس توصیه بیمه. سامانه‌های پردازشی و ارتباطی چندرسانه‌ای هوشمند.۳(۴)، ۱۳-۲۸.

بابااکبری، امیر؛ قره‌خانی، محسن؛ زین‌ساز، علی. (۱۴۰۱). *بررسی چالش‌ها و راهکارهای صنعت بیمه در دورن پساتحریم*. پژوهشکده بیمه، گروه پژوهشی عمومی بیمه. https://www.irc.ac.ir/fa-IR/Irc/4946/Articles/view/14643/1597/

پویا، علیرضا؛ فاضل ترشیزی، داود. (۱۳۹۹). بخش بندی مشتریان صنعت بیمه با استفاده از الگوریتم خوشه بندی دو مرحله ای مقیاس پذیر (نمونه موردی شرکت بیمه البرز). پژوهشنامه بیمه (صنعت بیمه)، ۳۵(۲)، ۳۱-۶۵.

پیران، فاطمه؛ علیرضا؛ بنی‌هاشمی، سیدعلی. (۱۴۰۲). تأثیر هوش تجاری و مؤلفه¬های آن بر ایجاد و توسعه یک سازمان چابک. *دوماهنامه‌های بررسی‌های بازرگانی*. ۲۱(۱۱۹)، ۱۱۹-۱۳۴.

تارخ، محمدجعفر؛ اکبری، تورج. (۱۳۹۴). استفاده از خوشه بندی در مدیریت ارتباط با مشتریان صنعت بیمه. *نخستین کنفرانس بین‌المللی فناوری اطلاعات*.

خلیلی دیلمی، سیدمرداویج ؛آقاسیدمیرزا، سیده شهره. (۱۴۰۳). نقش هوش تجاری در شرکت‌های بیمه. *مطالعات آینده پژوهی و سیاستگذاری*. ۱۰(۲)، ۱۰-۲۵.

رشیدی، محمدحافظ. (۱۴۰۲). بررسی سیستماتیک تاثیر و کاربردهای متنوع هوش تجاری در عملکرد سیستم‌ها. ششمین کنفرانس بین‌المللی مدیریت و صنعت (بین المللی). ۱۲۳-۱۳۲.

قاسمی، فهیمه سادات. (۱۳۹۷). شناسایی مشتریان مناسب از طریق جداول هوش تجاری توسط درخت تصمیم و نرم‌افزار رپید ماینر (مطالعه موردی بیمه پارسیان استان قم)، *دوماهنامه پژوهش‌های نوین علوم‌انسانی*، ۲ (۷)، ۱۹-۱.

سجودی شیجانی، امید؛ بوبه رژ، آتنا؛ و عبدالوند، ندا. (۱۳۹۳). اهمیت بندی مشخصه‌های RFM در صنعت بیمه. *همایش بین‌المللی بیمه و توسعه*.

سلیمانی، محمدرضا؛ عاطفت دوست، علیرضا. (۱۳۹۸). بررسی تأثیر هوش تجاری بر تصمیم‌گیری سازمانی با درنظرگرفتن نقش میانجی متغیرهای کیفیت داده و زمینه هوش تجاری (موردمطالعه: کارکنان فروشگاه کوثر قم)، *فصلنامه پژوهش‌های معاصر در علوم مدیریت و حسابداری*. ۱(۳)، ۶۹-۸۲.

محمدی، رضا. (۱۴۰۲)، هوش تجاری و انقلاب صنعتی چهارم: فرصت‌ها و چالش‌ها. نشریه تخصصی آرمان پردازش. ۴(۱)، ۱-۷.

محمد شریفی، محمد. (۱۳۹۸). ارزیابی ارزش مشتریان در حوزه بیمه بدنه اتومبیل با استفاده از روش‌های داده‌کاوی (مطالعه موردی)، (پایان‌نامه کارشناسی‌ارشد مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان).

مشفق، مصطفی؛ پارسا، عرفان؛ جلیلی صالح، نعمت.(۱۴۰۳). مدل ارزیابی عملکرد سازمانی با استفاده از ابزارهای هوش تجاری (BI): مطالعه موردی بانک ها. *مدیریت استراتژیک هوشمند*. ۳(۲)، ۷۳-۹۲.

مهدوی، غدیر؛ نصیری، فاطمه. (۱۳۹۷). اصول و مبانی نظری بیمه، *پژوهشکده بیمه*.

Aggarwal, A. G., & Yadav, S. (2020, June). Customer segmentation using fuzzy-AHP and RFM model. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 77-80). IEEE.

Aggarwal, C. C., & Reddy, C. K. (Eds.). (2014). Data clustering: Algorithms and applications. CRC Press.

Anitha, P., & Patil, M. M. (2022). RFM model for customer purchase behavior using K-Means algorithm. Journal of King Saud University-Computer and Information Sciences, 34(5), 1785-1792.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1), 1-27.

Baudry, M., & Robert, C. Y. (2019). A machine learning approach for individual claims reserving in insurance. Applied Stochastic Models in Business and Industry, 35(5), 1127-1155.

Christy, A. J., 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.

Das, S. (2024). Enhancing Agent Interactions and Decision-Making in Insurance with Intelligent Technologies. Journal of Computational Analysis & Applications, 33(8).

Davies, D. L., & Bouldin, D. W. (2009). A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, (2), 224-227.

Dehghanpour, S., & Esfahanipour, A. (2018). Dynamic portfolio insurance strategy: a robust machine learning approach. Journal of Information and Telecommunication, 2(4), 392-410.

Dogan, O., Seymen, O. F., & Hiziroglu, A. (2022). Customer behavior analysis by intuitionistic fuzzy segmentation: comparison of two major cities in Turkey. International Journal of Information Technology & Decision Making, 21(02), 707-727.

Ernawati, E., Baharin, S. S. K., & Kasmin, F. (2021, April). A review of data mining methods in RFM-based customer segmentation. In Journal of Physics: Conference Series (Vol. 1869, No. 1, p. 012085). IOP Publishing.

Ernst, C. P., & Geiger, F. (2021). Business Intelligence in the Database Marketing–A Case Study of a German Insurance Company

Gaikwad, D. D. (2020). Current Scenario of Different Type of Services Marketing and Delivery in Customer Relationship Management. International Journal of Engineering and Management Research,10.

Geetha, K. (2020). Param (2020). Data Analysis and ETL Tools in Business Intelligence. International Research Journal of Computer Science (IRJCS), 7, 127-131.

He, W., Wang, F. K., & Akula, V. (2017). Managing extracted knowledge from big social media data for business decision making. Journal of knowledge management, 21(2), 275-294.

Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. Expert Systems with Applications, 37(7),5259-5264.

Jamjoom, A. A. (2021). The use of knowledge extraction in predicting customer churn in B2B. Journal of Big Data, 8(1), 110.

John, J. M., Shobayo, O., & Ogunleye, B. (2023). An exploration of clustering algorithms for customer segmentation in the UK retail market. Analytics, 2(4), 809-823.

Kaur, R., & Gabrijelčič, D. (2022). Behavior segmentation of electricity consumption patterns: A cluster analytical approach. Knowledge-based systems, 251, 109236.

Khayatbashi, S., Sjölind, V., Granåker, A., & Jalali, A. (2025, June). AI-enhanced business process automation: a case study in the insurance domain using object-centric process mining. In International Conference on Business Process Modeling, Development and Support (pp. 3-18). Cham: Springer Nature Switzerland.

Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. Procedia computer science, 3, 57-63.

Kumar, N. (2025). Intelligent customer segmentation: unveiling consumer patterns with machine learning. Journal of Umm Al-Qura University for Engineering and Architecture, 1-10.

Lang, L., Zhou, S., Zhong, M., Sun, G., Pan, B., & Guo, P. (2023). A Big Data Based Dynamic Weight Approach for RFM Segmentation. Computers, Materials & Continua, 74(2).

Massaro, A., Vitti, V., Galiano, A., & Morelli, A. (2019). Business intelligence improved by data mining algorithms and big data systems: an overview of different tools applied in industrial research. Computer Science and Information Technology, 7(1), 1-21.

Miah, S. J., & Yeoh, W. (Eds.). (2018). Applying business intelligence initiatives in healthcare and organizational settings. IGI Global.

Moscoso-Zea, O., Castro, J., Paredes-Gualtor, J., & Luján-Mora, S. (2019). A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education. IEEE access, 7, 38778-38788.

Muradkhanli, L. G., & Karimov, Z. M. (2023). Customer behavior analysis using big data analytics and machine learning. Problems of Information Society, 61-67.

Nandapala, E. Y. L., Jayasena, K. P. N., & Rathnayaka, R. M. K. T. (2020, December). Behavior segmentation based micro-segmentation approach for health insurance industry. In 2020 2nd International Conference on Advancements in Computing (ICAC) (Vol. 1, pp. 333-338). IEEE.

Panigrahi, S., & Palkar, B. (2018). Comparative analysis on classification algorithms of auto-insurance fraud detection based on feature selection algorithms. Int. J. Comput. Sci. Eng, 6(9), 72-77.

Patil, K. S., & Godbole, A. (2018). A survey on machine learning techniques for insurance fraud prediction. Helix, 8(6), 4358-4363.

Pomffyová, M., & Bartková, L. (2016). Take advantage of information systems to increase competitiveness in SMEs. Procedia-Social and Behavioral Sciences, 220, 346-354.

Press, W. H. (2007). Numerical recipes 3rd edition: The art of scientific computing. Cambridge university press.

Rahmah, N., & Sitanggang, I. S. (2016). Determination of optimal epsilon (eps) value on dbscan algorithm to clustering data on peatland hotspots in sumatra. In IOP conference series: earth and environmental science (Vol. 31, No. 1, p. 012012). IoP Publishing.

Ramachandaran, S., Mahalley, Z., Nuraini, R., & Dhar, B. K. (2025). Exploring the challenges of AI-driven business intelligence systems in the Malaysian insurance industry. F1000Research, 14, 452.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, 53-65.

Sawadogo, P., & Darmont, J. (2021). On data lake architectures and metadata management. Journal of Intelligent Information Systems, 56(1), 97-120.

Sun, Z. H., Zuo, T. Y., Liang, D., Ming, X., Chen, Z., & Qiu, S. (2021). GPHC: A heuristic clustering method to customer segmentation. Applied Soft Computing, 111, 107677.

Turban, E., Sharda, R., Aronson, J. E., & King, D. (2008). Business intelligence: A managerial approach (pp. 58-59). Upper Saddle River, NJ: Pearson Prentice Hall.

Turban, E., Sharda, R., Aronson, J. E., & King, D. (2017). Business intelligence: A managerial approach (pp. 58-59). Upper Saddle River, NJ: Pearson Prentice Hall.

Wang, C. (2022). Efficient customer segmentation in digital marketing using deep learning with swarm intelligence approach. Information Processing & Management, 59(6), 103085.

Wit, E., Heuvel, E. V. D., & Romeijn, J. W. (2012). 'All models are wrong... ': an introduction to model uncertainty. Statistica Neerlandica, 66(3), 217-236.

Yuan, C., & Yang, H. (2019). Research on K-value selection method of K-means clustering algorithm. J, 2(2), 226-235.

Zhang, C., Wang, X., Cui, A. P., & Han, S. (2020). Linking big data analytical intelligence to customer relationship management performance. Industrial Marketing Management, 91, 483-494.

Zou, H. (2020). Clustering algorithm and its application in data mining. Wireless Personal Communications, 110(1), 21-30.