



ارائه چارچوبی به‌منظور بخش‌بندی بیمه‌گذاران بیمه عمر با استفاده از داده‌کاوی

مریم روستازاده شیخ یوسفی^۱

سید محمدرضا میراحمدی^۲

چکیده

هدف از این پژوهش ارائه چارچوبی برای بخش‌بندی بیمه‌گذاران بیمه عمر با استفاده از روش‌های داده‌کاوی است. جامعه آماری پژوهش شامل مشتریان شرکت بیمه‌ای در شهر اصفهان بوده است و داده‌های لازم از قراردادهای بیمه‌گذاران بیمه عمر این شرکت، در بازه زمانی سال‌های ۱۳۸۷ تا ۱۳۹۷ جمع‌آوری شده است. بدین ترتیب داده‌های مربوط به ۳۵۳ نفر از بیمه‌گذاران بیمه عمر در قالب چهارده متغیر (ویژگی‌های فردی و شرایط بیمه انتخابی) گردآوری شد. با استفاده از الگوریتم خوشه‌بندی کا - میانگین و به کمک معیار سیلوئت در نرم‌افزار «متلب»، مشتریان در چهار خوشه طبقه‌بندی شدند. تحلیل نتایج به‌دست‌آمده از رفتار مشتریان هر خوشه، مبنایی برای نام‌گذاری خوشه‌ها به نام‌های ترقی‌خواهان، محافظه‌کاران، زحمت‌کشان و طلایه‌داران به‌دست داد. همچنین به‌منظور تسهیل فرایند بررسی متغیرهای ورودی در هر خوشه، فقط هفت متغیر (سن، تحصیلات، شغل، نحوه پرداخت حق بیمه، تعداد استفاده‌کنندگان از سرمایه بیمه در صورت فوت، نرخ افزایش سرمایه فوت و پرداخت حق بیمه منظم) در تحلیل نهایی وارد شد، یعنی متغیرهایی که در خوشه‌هایی که در آزمون «خی» دو تفاوت معنادار بین آن‌ها در سطح ۰/۰۰۱ وجود داشت.

واژگان کلیدی: بخش‌بندی مشتریان، داده‌کاوی، الگوریتم کا - میانگین، بیمه عمر

تاریخ دریافت: ۱۳۹۸/۰۷/۰۷

تاریخ پذیرش: ۱۳۹۸/۱۰/۰۸

مقدمه

روانی و اجتماعی خلق می‌کند، بیمه‌های عمر و پس‌انداز است. فاصله زیاد بین دریافت حق بیمه و اجرای تعهدات شرکت‌های بیمه در بیمه‌نامه‌های عمر و پس‌انداز، که باعث تجمع ذخایر ریاضی درخور ملاحظه‌ای در ارتباط با هر بیمه‌نامه می‌شود، به شرکت‌های بیمه این امکان را می‌دهد که با استفاده بهینه از آن، بازار سرمایه را فعال کنند و نقش جدی‌تری در فعالیت‌های سرمایه‌گذاری داشته

وجود ظرفیت‌های بالقوه اقتصادی فراوان در ایران، که هنوز به فعلیت نرسیده است یا با همه ظرفیت مشغول به کار نیست، از جمله معضلات بزرگ مدیریتی کشور به‌شمار می‌رود. صنعت بیمه^۱ کشور، به‌ویژه بیمه عمر^۲ و پس‌انداز نیز از این قاعده مستثنا نیست (ترکی هرچگانی، ۱۳۹۴). یکی از انواع بیمه، که محیط امنی را برای فعالیت‌های اقتصادی و سرمایه‌گذاری و امنیت روحی

۱. کارشناس ارشد مدیریت بازرگانی، گروه مدیریت، واحد مبارکه، دانشگاه آزاد اسلامی.

۲. استادیار گروه مدیریت، واحد دولت‌آباد، دانشگاه آزاد اسلامی، اصفهان (نویسنده مسئول)؛ M.mirahmadi@gmail.com

3. Life Insurance

4. Insurance Industry

یعنی جلب اعتماد مشتری و سودآوری بالا در شرکت بیمه دست یافت. پرسش اساسی پژوهش حاضر این است که چگونه می‌توان چارچوبی را به منظور بخش‌بندی بیمه‌گذاران عمر، با توجه به روش داده‌کاوی، ارائه کرد؟

۱. مبانی نظری پژوهش

۱-۱. بیمه عمر

بی‌شک بیمه در بهبود کیفیت زندگی افراد جامعه نقش بسزایی دارد و توسعه تأمین اجتماعی در این‌باره اهمیت بالایی دارد (Kaigorodova and Mustafina, 2014). بیمه به معنای تعاون و همیاری اجتماعی به منظور سرشکن کردن زیان افراد جامعه است (کریمی، ۱۳۹۰). به طور کلی، خدمات بیمه به سه دسته تقسیم می‌شوند: بیمه مسئولیت، بیمه اموال و بیمه اشخاص. بیمه مسئولیت بخشی از سیستم بیمه‌های عمومی است که مستقیماً در تنظیم روابط اجتماعی افراد جامعه، شناخت افراد از حقوق و مسئولیت‌های یکدیگر و همچنین تأمین امنیت حرفه‌ای مشاغل تأثیرگذار است؛ بیمه مسئولیت مدنی یا مسئولیت پزشک از این دست بیمه‌ها به‌شمار می‌رود (طاهری، ۱۳۹۷). بیمه‌های اموال شامل بیمه آتش‌سوزی، بیمه باربری، بیمه اتومبیل، بیمه مهندسی، بیمه هواپیما، بیمه انفجار اتمی، بیمه تعطیلی کارخانه‌ها و بیمه سرقت^۳ و بیمه‌های اشخاص نیز شامل بیمه‌های زندگی (عمر و پس‌انداز)، بیمه حوادث و بیمه درمانی است (بیمه مرکزی ایران، ۱۳۹۰ به نقل از عابدی، ۱۳۹۵). بیمه عمر قراردادی است که به‌موجب آن، بیمه‌گر در مقابل دریافت حق بیمه متعهد می‌شود که در صورت فوت بیمه‌شده یا در صورت زنده ماندن بیمه‌شده در موعد مقرر در قرارداد، مبلغ بیمه تعیین شده را به بیمه‌گذار (یا بیمه‌شده) یا استفاده‌کننده معین بپردازد (حلاج‌نژادی، ۱۳۹۷). در این بیمه، افراد با پرداخت مبلغی به‌صورت ماهیانه یا سالانه با عنوان حق بیمه، افزون‌بر تأمین منابع لازم برای تشکیل سرمایه‌ای در آینده برای خود، از امکانات پوشش‌های بیمه عمر نیز بهره‌مند خواهند شد. به عبارتی، اگر حادثه و اتفاق ناگواری مانند فوت، نقض عضو، ازکارافتادگی یا بیماری‌های خاص اتفاق بیفتد، بیمه‌حاملی افراد خواهد بود و نه فقط مبلغ سرمایه‌گذاری شده و سود آن را پرداخت می‌کند، بلکه مبلغی هم در حکم پوشش بیمه‌ای می‌پردازد تا مشکلات مالی را هموار کند (ترکی هرچگانی، ۱۳۹۵). بر این اساس، تقسیم‌بندی گوناگون بیمه‌های عمر و پس‌انداز عبارت است از بیمه‌های تمام عمر، بیمه عمر زمانی، بیمه‌های مختلط پس‌انداز^۴، بیمه‌های مستمری^۵، بیمه‌های بازنشستگی^۶ و بیمه‌های عمر گروهی^۷. هدف اساسی بیمه‌های عمر، تشکیل و جمع‌آوری ذخایر مالی،

باشند و به نوبه خود باعث افزایش اشتغال و رشد پایدار اقتصادی و برقراری امنیت اجتماعی برای جامعه و رونق بازارهای مالی شوند (عباسی و حسینی، ۱۳۹۷).

سازمانی که تصمیم می‌گیرد در بازاری گسترده فعالیت کند، می‌داند که معمولاً از عهده ارائه خدمات به همه مشتریان آن بازار برنمی‌آید. بازاریابی هدفمند شامل شناسایی سودآورترین بخش‌های بازار است؛ بنابراین مشاغل ممکن است تصمیم بگیرند فقط روی یک یا چند مورد از این بخش‌ها تمرکز کنند. آن‌ها ممکن است محصولات یا خدماتشان را برای برآورده کردن هر بخش انتخاب‌شده توسعه دهند. این راهبرد بازاریابی هدفمند با بازاریابی گسترده یا راهبرد تمایز محصول متفاوت است (Camilleri, 2017). انتخاب راهبرد بازاریابی عاقلانه نیازمند درک دقیق رفتار مصرف‌کننده است که لازمه برآورده کردن بهتر نیازهای مشتریان و شناخت و تجزیه و تحلیل صحیح آنان است، امری که فقط از راه بخش‌بندی بازار^۱ آن‌ها بهتر میسر می‌شود (شیرخدایی و همکاران، ۱۳۹۵). بخش‌بندی بازار یکی از نخستین مراحل حیاتی مدیریت ارتباط با مشتریان است؛ زیرا به دسته‌بندی مشتریان براساس نوع خواسته‌های متفاوت آن‌ها در دنیای پر تغییر امروز می‌پردازد (Chien-Wen, 2012). از آنجاکه روش‌های سنتی جواب‌گوی نیازهای جدید بازار نیست، یکی از جدیدترین روش‌های بخش‌بندی، استفاده از رویکرد داده‌کاوی^۲ است که در قیاس با الگو و روش‌های سنتی، در پیچه جدیدی به شرکت‌ها می‌دهد (Goetzke and Spiller, 2014). مشخصات کلیدی داده‌کاوی شامل کشف خودکار الگوها، پیش‌بینی خروجی‌های محتمل، ساخت اطلاعات قابل استناد، تمرکز بر مجموعه‌های داده‌ای و پایگاه داده‌های بزرگ است (Destpand and Thaker, 2010).

به‌رغم آنکه شرکت‌های بیمه عمر برای بهبود کارایی تجاری خود از روش‌های نوین تجزیه و تحلیل استفاده می‌کنند، اما هنوز خلاء ارائه مطالعات گسترده درباره اینکه چگونه تجزیه و تحلیل‌های پیش‌بینی‌کننده می‌تواند حوزه بیمه عمر را غنی کند حس می‌شود (Boodhun and Jayabalan, 2018)؛ بنابراین ضعف در شناسایی بازار، فقدان طراحی محصول منطبق با نیازهای بیمه‌گذار، فقدان بازار رقابتی در حوزه بازاریابی و فروش و فقدان نیروی متخصص برای ارزیابی ریسک از جمله مواردی است که تأثیر نامناسبی در ارائه خدمات به بیمه‌گذاران می‌گذارد (رجوعی و شیعه‌زاده، ۱۳۸۷). با استفاده از نتایج این تحقیق، می‌توان به شناسایی نیازهای اساسی بیمه‌گذار و نقاط ضعف در ارائه خدمات پرداخت و نحوه رفتار و نوع محصول خریداری شده به‌دست بیمه‌گذاران را پیش‌بینی کرد تا به دو هدف عمده در صنعت بیمه،

3. Insurance of Theft

4. Mixed Savings Insurance

5. Pension Insurance

6. Retirement Insurance

7. Group Life Insurance

1. Market Segmentation

2. Data Mining

جدول ۱: معیارهای بخش‌بندی اثربخش (Bingham et al., 2005)

مفهوم	معیار
بخش‌های گوناگون شناسایی شده قابلیت ارزیابی دارند.	۱. قابلیت ارزیابی
همگونی درون بخش‌ها و ناهمگونی بین بخش‌ها	۲. قابلیت تمییز
بخش‌های هدف توجیه‌کننده رفتار خاص شرکت باشند.	۳. اهمیت
بخش‌های هدف پاسخ‌گوی توسعه آمیخته بازاریابی برای هریک از بخش‌ها باشند.	۴. عملکردی بودن
قابلیت دسترسی به هریک از بخش‌ها از راه ارتباطات بازاریابی	۵. دست‌یافتنی بودن

تقسیم بازار باعث می‌شود بازار نامتجانس به بخش‌هایی با خصوصیات متجانس تقسیم شود. با این کار، هر سازمان می‌تواند در هر بخش از بازار، همه نیازهای خریدارانش را برآورده کند و از این راه، میزان فروش و سود خود را بالا ببرد (بهروزیان، ۱۳۹۵). مک دونالد و دونبار (2012) پنج مورد از مزایای بخش‌بندی را برشمرده‌اند که عبارت‌اند از:

۱. برای موفقیت بازاریابی، تشخیص تفاوت‌های مشتریان امری کلیدی است که می‌تواند نزدیک‌شدن محصولات و خدمات شرکت به نیازهای مشتریان را فراهم کند؛

۲. بخش‌بندی می‌تواند در جای مناسب به بازاریابی دقیق و جزئی منجر شود، جایی که شرکت می‌تواند در نتیجه بخش‌بندی جزئی در بخش‌های تشکیل‌شده نیازهای مشتریان را برآورده کند، چیزی که اغلب در بازاریابی کلی نیز غیرممکن است؛

۳. بخش‌بندی می‌تواند در بازارهایی که مزیت رقابتی بیشتر و درآمدها بالاست منابع را متمرکز کند؛

۴. بخش‌بندی می‌تواند برای رسیدن به مزیت رقابتی استفاده شود؛ به این صورت که شرکت را قادر می‌سازد، به نسبت رقبایش، بازار را از ابعاد گوناگون در نظر بگیرد؛

۵. با بخش‌بندی بازار می‌توان با درک بهتر از نیازهای مشتریان و خلق مزایای بیشتر در محصولات و خدمات شرکت، در قیاس با رقبای، به صورت موثرتری بازاریابی کرد.

۱-۳. داده داده‌کاوی

فناوری‌های تحلیل داده به منظور جمع‌آوری، پردازش، تجزیه و تحلیل، و مدیریت مؤثر داده‌ها روش‌های نوینی را برای شرکت‌های بیمه طراحی می‌کنند؛ بنابراین در بخش‌های گوناگونی از صنایع بیمه، مانند ارزیابی ریسک، تجزیه و تحلیل رفتار مشتری، توسعه محصول و تحلیل بازاریابی استفاده می‌شوند (Boothun and Jayabalan, 2018). اصطلاح کشف دانش برای نخستین بار در دهه ۱۹۹۰ مطرح شد و توجه پژوهشگران را به الگوهای داده‌کاوی معطوف کرد. هدف داده‌کاوی کشف دانش

سرمایه‌گذاری و کسب سود حاصل از سرمایه‌گذاری و در نهایت ایفای تعهدات شرکت بیمه‌گر در قبال استفاده‌کنندگان است. بیمه‌های عمر و ذخایر انباشته آن به‌منزله عامل در ارتقای سطح تولید ملی نقش سازنده ایفا می‌کند و این افزایش سطح تولید پرداخت سرمایه‌ها و غرامت‌های کلیه رشته‌های اشخاص در تسریع گردش چرخ‌های اقتصادی کشور مؤثر خواهد بود (شادکام اصفهانی، ۱۳۹۵).

۱-۲. بخش‌بندی بازار

در فضای رقابتی امروز، هر شرکتی که بخواهد در هر بازاری، اعم از بازارهای صنعتی و بازارهای مصرفی، فعالیت کند باید بدانند که نمی‌تواند همه سلاقی و درخواست‌های مشتریان را یک‌جا پاسخ دهد و به همه آن‌ها خدمت‌رسانی کند (اکبری و همکاران، ۱۳۹۵). لازمه برآورده کردن بهتر نیازهای مشتریان و ارائه ارزش بیشتر به آن‌ها، شناخت و تجزیه و تحلیل صحیح مشتریان است، امری که فقط از راه بخش‌بندی آن‌ها بهتر میسر می‌شود (شیرخدایی و همکاران، ۱۳۹۵). بخش‌بندی بازار عبارت است از فرایند تقسیم بازارها به زیرمجموعه‌های مشخص که در آن‌ها، مشتریان رفتارها و نیازهای یکسانی دارند و هریک از این زیرمجموعه‌ها را می‌توان به‌منزله بازار هدفی قلمداد کرد که با راهبرد (راهبرد)‌های مشخص می‌توان به آن‌ها دسترسی یافت (Trindade et al., 2017). در واقع بخش‌بندی بازار زیرمجموعه‌ای از افراد یا سازمان‌های یک بازار با یک یا چند ویژگی مشابه است که باعث می‌شود آن‌ها تقاضاهای مشابهی از محصولات و خدمات مبتنی بر ویژگی‌های آن محصولات، مانند قیمت و کاربردش، داشته باشند (Mahalakshmi and Chitra, 2014).

بخش‌بندی بازار به سازمان امکان می‌دهد که فعالیت‌های کلیدی بازاریابی خود را برای هدف قراردادن گروه خاصی از مشتریان متمرکز کند. این گروه خاص از بیشترین احتمال برای خرید محصولات و خدمات سازمان بهره‌مند بوده است یا مناسب‌ترین گروه برای بهره‌برداری از محصولات و خدمات سازمان بوده است (Lamberti and Noci, 2009). معیارهای بخش‌بندی اثربخش را می‌توان در پنج مورد مطابق جدول ۱ در نظر گرفت:

1. Segmentation

جدید، معتبر و پیگیری‌شونده با استفاده از ابزارهای هوش مصنوعی و آماری در حجم بالای داده است (Marbin et al., 2012). داده‌کاوی یا کشف دانش در پایگاه داده‌ها ابزاری قدرتمند است برای استخراج دانش بالقوه نهفته و اطلاعات پیشین سودمند که از مجموعه‌ای از داده‌ها به‌کار می‌رود. این فرایند به‌صورت خودکار به کشف روابط و الگوهای موجود در داده‌های خام و اجرای نتایج آن می‌پردازد (شاهینی و همکاران، ۱۳۹۵). هدف اصلی داده‌کاوی پیش‌بینی است و به‌گونه‌ای دقیق‌تر می‌توان گفت داده‌کاوی شناسایی الگوهای صحیح، بدیع، سودمند و درک‌شدنی از داده‌های موجود در هر پایگاه داده است که با استفاده از پردازش‌های معمول نمی‌توان به آن‌ها دسترسی یافت (Hand et al., 2011). چرخه عمر هر پروژه داده‌کاوی شامل شش مرحله است: درک مسئله، کسب و کار، درک داده‌ها، آماده‌سازی داده، مدل‌سازی، ارزیابی نتایج و به‌کارگیری مدل (آذر و همکاران، ۱۳۸۹).

$$J = \sum_{j=1}^k \sum_{i=1}^n \|X_i^{(j)} - C_j\|^2$$

|| معیار فاصله بین نقاط و C_j مرکز خوشه j ام.

الگوریتم زیر الگوریتم پایه برای این روش به‌شمار می‌رود:

۱. نخست k نقطه به صورت تصادفی به عنوان مراکز خوشه‌ها انتخاب می‌شوند؛

۲. هر نمونه داده به خوشه‌ای نسبت داده می‌شود که مراکز آن خوشه کمترین فاصله تا آن داده را داراست؛

۳. پس از تعلق تمامی داده‌ها به یکی از خوشه‌ها، برای هر خوشه یک نقطه جدید به‌منزله مرکز محاسبه می‌شود (میانگین نقاط متعلق به هر خوشه).

مراحل ۲ و ۳ تکرار می‌شوند تا زمانی که دیگر هیچ تغییری در مراکز خوشه‌ها حاصل نشود (Celby et al., 2013). در روش کا - میانگین با تعیین مراکز خوشه، تعداد خوشه‌ها از قبل تعیین شده‌اند و به تخصیص مشاهدات به هر خوشه با هدف به حداقل رساندن فاصله هر مشاهده از مرکز خوشه اقدام می‌کند. تکرار این الگوریتم تا زمان فقدان تغییر در مراکز خوشه‌ها در تکرار متوالی ادامه می‌یابد. در این تکنیک، الگوها و روابط پنهان میان مجموعه داده با به حداقل رسیدن فواصل درون خوشه معین می‌شود (مؤمنی، ۱۳۹۰).

۱-۴. پیشینه پژوهشی

باش افشار و همکاران (۱۳۹۷) در پژوهشی به بررسی اطلاعات مشتریان بیمه عمر به‌منظور دستیابی به الگوی خوشه‌بندی برای ارائه خدمات پرداختند. از بین شرکت‌های بیمه عمر، شرکت بیمه‌ای با اندازه نمونه ۱۰۰۰ نفر از مشتریان، که به خرید بیمه‌نامه عمر اقدام کرده بودند، انتخاب شد. نتایج حاصل از داده‌کاوی نشان داد که مشتریان بیمه عمر را می‌توان در دو گروه اصلی «مشتریان سودبخش» و «مشتریان دارای ریسک» طبقه‌بندی کرد؛ همچنین متغیرهای جمعیتی «جنسیت» و «سن» و متغیرهای بیمه‌ای همچون «حق بیمه سالانه» و «ضریب فوت بر اثر حادثه»، در حکم عوامل تأثیرگذار در گروه‌های مشتریان شناسایی شد.

بهناز و حسینی (۱۳۹۷) پژوهشی را با هدف دسته‌بندی سرویس‌های مشتریان به لحاظ سطح استفاده از خدمات شبکه شتاب با استفاده از مدل‌های ترکیبی ارائه دادند. در این تحقیق برای شناسایی خدمات سودآوری که مشتریان به‌خوبی از آن استقبال کردند از چهار روش داده‌کاوی درخت تصمیم، بیز ساده، K همسایه نزدیک

تکنیک‌های گوناگون داده‌کاوی را می‌توان براساس نوع عملیاتی که ارائه می‌دهند به دو دسته پیش‌بینی‌کننده و تشریح‌کننده تقسیم کرد. تکنیک‌های پیش‌بینی‌کننده با ساخت مدلی برای پایگاه داده، وظیفه پیش‌بینی موارد ناشناخته را برعهده دارند؛ درحالی‌که تکنیک‌های تشریح‌کننده الگوهایی فهمیدنی از داده را برای انسان کشف می‌کنند (Blonder et al., 2019). خوشه‌بندی از مفیدترین کارکردهای داده‌کاوی برای کشف گروه‌ها و تعیین توزیع‌های مدنظر و الگوها در داده‌هاست. مسئله خوشه‌بندی در مورد جداسازی هر مجموعه داده به گروه‌ها به‌نحوی است که داده‌های موجود در یک خوشه به‌نسبت نقاط موجود در خوشه دیگر شباهت بیشتری به یکدیگر داشته باشند؛ برای مثال بخش‌بندی بیمه‌گذاران فعلی به گروه‌هایی مشخص و مرتبط کردن یک پروفایل با هر گروه می‌تواند در راهبردهای نرخ‌گذاری آینده مفید باشد (قره‌خانی و ابوالقاسمی، ۱۳۹۴). در این تکنیک، یک پایگاه داده n شی به k گروه بخش‌بندی می‌شود. این کار با معیاری انجام داده می‌شود که برای گروه‌بندی در نظر گرفته شده است. روش‌های گوناگونی برای دسته‌بندی ارائه شده است که روش کا - میانگین از جمله مهمترین روشهای خوشه‌بندی است (عاملی بصیری و قره‌خانی، ۱۳۹۳).

۱-۳-۱. خوشه‌بندی کا - میانگین

در نوع ساده‌ای از روش خوشه‌بندی کا - میانگین،^۱ نخست به تعداد خوشه‌های لازم نقاطی به‌صورت تصادفی انتخاب می‌شوند. سپس، با توجه به میزان نزدیکی یا شباهت، به یکی از این خوشه‌ها نسبت داده می‌شود و خوشه‌های جدیدی به‌دست می‌آید. این مراحل تکرار می‌شوند. می‌توان در هر تکرار با میانگین‌گیری از داده‌ها، مراکز جدیدی برای آن‌ها محاسبه کرد و دوباره داده‌ها را

1. K-Means Clustering

را به شکل مؤثر بسنجد. مقایسه بین مدل‌های تحلیلی نشان داد که روش جنگل تصادفی به نسبت سایر روش‌های پیش‌بینی سنتی مانند رگرسیون خطی، درخت تصمیم و مدل تقویت تصمیم‌یافته عملکرد بهتری دارد. مطالعه تجربی نشان داد که منطقه مشتری، سن، وضعیت بیمه، جنس و منبع مشتری مهم‌ترین عوامل برای پیش‌بینی سودآوری مشتری بیمه است.

۲. روش‌شناسی پژوهش

این پژوهش از نظر هدف کاربردی است و به علت استفاده از اطلاعات قراردادهای بیمه عمر برای خوشه‌بندی در زمره تحقیقات پیمایشی قرار دارد. جامعه آماری پژوهش شامل مشتریان بیمه عمر پارسیان در نظر گرفته شد و اطلاعات از قراردادهای بیمه‌گذاران بیمه عمر این شرکت در شهر اصفهان در بازه زمانی سال‌های ۱۳۸۷ تا ۱۳۹۷ جمع‌آوری شد. با فرض بالابودن حجم جامعه، برای برآورد حداکثر مشابهت نمونه با جامعه، مطابق جدول مورگان، حجم نمونه برابر با ۳۸۴ نفر در نظر گرفته شد. با توجه به محدودیت‌های موجود در دسترسی و گردآوری اطلاعات و نواقص موجود در قراردادهای عقدشده با مشتریان، درنهایت داده‌های حاصل از قرارداد ۳۵۳ نفر از بیمه‌گذاران بیمه عمر، که به صورت کامل در سیستم ثبت شده بود و تمامی اطلاعات مدنظر پژوهشگر را دربر داشت، وارد فرایند نهایی تحلیل شد. این اطلاعات شامل هفت ویژگی فردی بیمه‌گذار (سن، جنسیت، وضعیت تأهل، میزان تحصیلات، وضعیت شغلی، محل سکونت، نسبت بیمه‌گذار با بیمه‌شده) و هفت ویژگی بیمه انتخابی بیمه‌گذار (نحوه پرداخت حق بیمه، پوشش اضافی، تعداد استفاده‌کنندگان از سرمایه بیمه در صورت حیات، تعداد استفاده‌کنندگان از سرمایه بیمه در صورت فوت، نسبت با بیمه‌شده در صورت فوت، نرخ افزایش سرمایه در صورت فوت، نرخ افزایش حق بیمه منظم) بود. در تجزیه و تحلیل داده‌ها و اطلاعات و برای ارائه چارچوبی با هدف بخش‌بندی بیمه‌گذاران بیمه عمر از روش داده‌کاوی و تکنیک الگوریتم ک- میانگین با معیار سنجش سیلوئت در نرم‌افزار متلب^۱ استفاده شد. گفتنی است که در پژوهش حاضر برای گردآوری اطلاعات از ابزار یا مقیاس به صورت پرسش‌نامه‌ای استفاده نشد و داده‌های پژوهش براساس وضعیت قراردادهای ثبت شده در شرکت بیمه پارسیان در بازه زمانی ده‌ساله گردآوری شد؛ بنابراین نیازی به بررسی روایی مقیاس نبود. همچنین برای ارزیابی اعتبار خوشه‌بندی و تصمیم‌گیری درباره انتخاب خوشه‌های بهینه، از میانگین شاخص سیلوئت استفاده شد.

۳. تجزیه و تحلیل داده‌های پژوهش

برای شروع خوشه‌بندی به روش ک- میانگین باید تعداد خوشه‌ها مشخص باشد. با توجه به حجم داده‌های در اختیار و بررسی

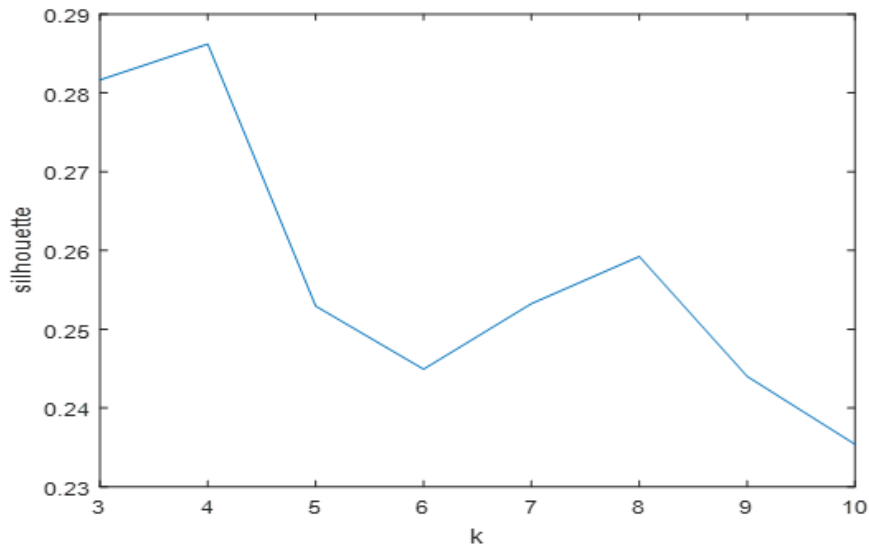
و مدل ترکیبی استفاده شد. نتایج ارزیابی مدل‌ها دقت مدل K همسایه نزدیک در شناسایی خدمات سودآور ۹۳/۲۶ درصد، دقت مدل بیز ۷۴/۸۳ درصد و دقت مدل درخت تصمیم ۹۷/۱۸ درصد است و پس از ترکیب مدل‌ها، دقت روش ترکیبی ۹۴/۸۰ درصد را نشان داد. همچنین داده‌های به دست آمده به چهار گروه پلاتینیومی، طلایی، نقره‌ای و سربی دسته‌بندی شده است. از مقایسه نتایج ارزیابی مدل‌ها مشخص شد که استفاده از روش‌های ترکیبی به نسبت روش‌های دسته‌بندی مبتنی بر یک مدل، از دقت بیشتری برخوردار است.

شریفی (۱۳۹۶) پژوهشی را با هدف دسته‌بندی مشتریان و تحلیل آن‌ها به منظور تعیین سودآورترین و ضرررسان‌ترین نوع بیمه برای شعبه بیمه و شرکت بیمه براساس الگوریتم‌های درخت تصمیم، K-medoids و شبکه بیزین ارائه داد. با اعتبارسنجی متقابل K-fold مجموعه داده ارزیابی شد و برای هر دسته، میزان قدر مطلق خطا محاسبه شد همچنین، با توجه به نتایج این الگوریتم‌ها، الگوریتم درخت تصمیم بهترین نتیجه را به نسبت دیگر روش‌ها داشته است. با توجه به نتایج تجزیه و تحلیل کلاس‌ها، گفتنی است به ترتیب کلاس‌های F (مهندسی)، E (مسئولیت)، A (عمر و اندوخته‌ساز)، B (آتش‌سوزی)، G (درمان) و درنهایت کلاس C (اتومبیل) بیشترین سود را برای این شعبه شرکت بیمه داشته است.

عبدی و همکاران (2018) پژوهشی را با هدف ارائه رویکردی سه‌مرحله‌ای داده‌کاوی برای تشخیص مشتریان با قابلیت اعتماد بالا و پیش‌بینی میزان فروش پوشش‌های ویژه شرکت بیمه ایران ارائه دادند. مرحله اول شامل غربالگری داده‌ها، مرحله دوم شامل انتخاب پوشش‌ها و ویژگی‌های مناسب و مرحله سوم استفاده از الگوریتم ک- میانگین به منظور دسته‌بندی مشتریان بود. هدف از این رویکرد، انتخاب زیرمجموعه گروهی فشرده با حداکثر قابلیت پیش‌بینی است. رویکرد پیشنهادی می‌تواند مشتریانی را شناسایی کند که بیشتر احتمال خرید پوشش بیمه‌ای خاص را در پایان دوره قرارداد دارند. نتایج نشان داد که رویکرد پیشنهادی قادر نیست خوشه‌های مشتری را تشخیص دهد و برای یک برنامه فروش پوشش بیمه مناسب برای مشتریان وفادار، با سطح دقت مناسب برنامه‌ریزی کند.

وو و همکاران (2017) پژوهشی را با هدف یک الگوریتم جنگل تصادفی برای تجزیه و تحلیل داده‌های بزرگ بیمه ارائه دادند. برای گردآوری اطلاعات، داده‌های مشتریان از شرکت‌های بیمه عمر چین جمع‌آوری شد تا نحوه رفتار مشتریان را با استفاده از الگوریتم پیشنهادی تحلیل شود. نتایج آزمایش‌ها نشان داد که الگوریتم جنگل تصادفی بهتر از SVM و رگرسیون لجستیک در عملکرد و دقت در داده‌های نامتوازن عمل می‌کند.

فنگ و همکاران (2017) پژوهشی را با هدف پیش‌بینی سودآوری مشتری با استفاده از تجزیه و تحلیل داده‌های بزرگ در صنعت بیمه ارائه دادند. با توجه به پیشینه رفتار خرید و ویژگی‌های فردی مشتریان، روش پیشنهادی می‌تواند سهم واقعی مشتری بیمه



نمودار ۱: تعیین تعداد بهینه خوشه‌ها با استفاده از شاخص سیلوئت

با بررسی مقادیر متغیرهای ورودی، به هر خوشه نامی اختصاص داده شده است و علت نام‌گذاری هر خوشه در نتیجه‌گیری بررسی شده است. این نام‌ها عبارت‌اند از ترقی‌خواهان برای خوشه ۱، محافظه‌کاران برای خوشه ۲، زحمت‌کشان برای خوشه ۳ و طلایه‌داران برای خوشه ۴.

۳-۱. توصیف خوشه‌ها با توجه به متغیرهای منتخب

برای ارزیابی متغیرهای ورودی در هر خوشه از جداول توافقی در نرم‌افزار SPSS-22 استفاده شده است. از آنجاکه متغیرهای ورودی پژوهش شامل چهارده متغیر بود، برای تسهیل فرایند بررسی متغیرهای ورودی در هر خوشه، تنها متغیرهایی وارد تحلیل نهایی شدند که در آزمون خی دو تفاوت معنادار بین آنها در سطح ۰/۰۰۱ وجود داشت. با توجه به این موضوع، متغیرهای سن ($\chi^2 = 33/66; p=0/001$)، تحصیلات ($\chi^2 = 193/13$); $p=0/001$)، شغل ($\chi^2 = 112/08; p=0/001$)، نحوه پرداخت حق بیمه ($\chi^2 = 25/59; p = 0/001$)، تعداد استفاده‌کنندگان از سرمایه بیمه در صورت فوت ($\chi^2 = 24/25; p=0/001$)، نرخ افزایش سرمایه فوت ($\chi^2 = 71/19; p = 0/001$) و پرداخت حق بیمه منظم ($\chi^2 = 65/21; p = 0/001$) در بین چهار خوشه

پژوهش‌های مشابه (برای مثال Trindade et al., 2017; Amorim and Komisarczuk, 2012; Celebi et al., 2013) خوشه در نظر گرفته شد و تعداد خوشه‌ها در هر مرحله یک واحد افزایش داده شده و معیار سیلوئت برای هر مرحله، محاسبه شد و این روند اجرا و ارزیابی در نرم‌افزار متلب انجام داده شد.

نمودار ۱ نشان می‌دهد که بهترین نقطه برازش بر روی داده‌ها در چهارمین خوشه (با ضریب برابر با ۰/۲۸۶+) بوده است؛ زیرا پس از آن و در خوشه پنجم، شیب نمودار با شدت به سمت پایین تمایل یافته است. ضریب به دست آمده بدان معناست که خوشه‌بندی نمونه به خوبی صورت گرفته است و چهار خوشه پیشنهاد شده برای نمونه مدنظر مناسب بوده است.

نتایج خوشه‌بندی به تعداد چهار خوشه نشان می‌دهد که در خوشه اول فراوانی برابر ۸۶ نفر و معادل با ۲۴ درصد بیمه‌گذاران، در خوشه دوم فراوانی برابر ۷۷ نفر و معادل با ۲۱ درصد بیمه‌گذاران، در خوشه سوم فراوانی برابر ۱۱۱ نفر و معادل با ۳۱ درصد بیمه‌گذاران و در خوشه چهارم فراوانی برابر ۷۹ نفر و معادل با ۲۲ درصد بیمه‌گذاران است. تابع هدف برابر با ۱۱۴/۲۱ و معیار سیلوئت برابر با ۰/۲۸۶ است. این نتایج در جدول ۲ به نمایش درآمده است.

جدول ۲: شاخص‌های فراوانی و درصد فراوانی در هر خوشه

الگوریتم	خوشه	فراوانی	درصد	تابع هدف	معیار سیلوئت
کا- میانگین	۱	۸۶	۲۴/۳۶	۱۱۴/۲۱	۰/۲۸۶
	۲	۷۷	۲۱/۸۱		
	۳	۱۱۱	۳۱/۴۴		
	۴	۷۹	۲۲/۳۸		

جدول ۳: جدول توافقی متغیر سن به تفکیک هر خوشه

سن				فراوانی (درصد)	خوشه		
بیش از ۵۰ سال	۴۱ تا ۵۰ سال	۳۱ تا ۴۰ سال	زیر ۳۰ سال				
۸ (۹/۳٪)	۱۸ (۲۰/۹٪)	۵۱ (۵۹/۳٪)	۹ (۱۰/۵٪)			۱. ترقی خواهان	
۲۶ (۳۳/۸٪)	۱۹ (۲۴/۷٪)	۲۳ (۲۹/۹٪)	۹ (۱۱/۷٪)				۲. محافظه کاران
۲۲ (۱۹/۸٪)	۴۶ (۴۱/۱٪)	۳۷ (۳۳/۳٪)	۶ (۵/۴٪)				۳. زحمت کشان
۱۵ (۱۹٪)	۲۵ (۳۱/۶٪)	۳۱ (۳۹/۲٪)	۸ (۱۰/۱٪)	۴. طلایه داران			

بیشترین فراوانی را دارند. نتایج جدول ۴ نشان می‌دهد که در رده تحصیلی زیر دیپلم و دیپلم، بیشترین فراوانی مربوط به خوشه زحمت کشان، در رده تحصیلی کاردانی، دو خوشه ترقی خواهان و محافظه کاران و در رده تحصیلی کارشناسی و تحصیلات تکمیلی، خوشه طلایه داران بیشترین فراوانی را دارند. نتایج جدول ۵ نشان می‌دهد که در مشاغل کارگری ساده و

تفاوت معناداری در سطح ۰/۰۰۱ داشتند. در ادامه، جداول توافقی مربوط به هر خوشه در هفت متغیر انتخاب شده بررسی شده است که نتایج آن در جداول شماره ۳ تا ۹ ارائه شده است. نتایج جدول ۳ نشان می‌دهد که در رده سنی زیر ۳۰ سال، بیشترین فراوانی مربوط به خوشه محافظه کاران، در رده سنی ۳۱ تا ۴۰ سال، خوشه ترقی خواهان، در رده سنی ۴۱ تا ۵۰ سال، خوشه زحمت کشان و در رده سنی بیش از ۵۰ سال، خوشه محافظه کاران

جدول ۴: جدول توافقی متغیر تحصیلات به تفکیک هر خوشه

تحصیلات					فراوانی (درصد)	خوشه		
تکمیلی	کارشناسی	کاردانی	دیپلم	زیر دیپلم				
۸ (۹/۳٪)	۲۶ (۳۰/۲٪)	۱۸ (۲۰/۹٪)	۲۵ (۲۹/۱٪)	۹ (۱۰/۵٪)			۱. ترقی خواهان	
۳ (۳/۹٪)	۲۵ (۳۲/۵٪)	۱۸ (۲۳/۴٪)	۲۷ (۳۵/۱٪)	۴ (۵/۲٪)				۲. محافظه کاران
۰ (۰٪)	۰ (۰٪)	۶ (۵/۴٪)	۵۹ (۵۳/۲٪)	۴۶ (۴۱/۴٪)				۳. زحمت کشان
۱۷ (۲۱/۵٪)	۴۴ (۵۵/۷٪)	۱۳ (۱۶/۵٪)	۵ (۶/۳٪)	۰ (۰٪)	۴. طلایه داران			

جدول ۵: جدول توافقی متغیر شغل به تفکیک هر خوشه

شغل				فراوانی (درصد)	خوشه		
صاحب بنگاه	کارمند	کارگر ماهر	کارگر ساده				
۱۶ (۱۸/۶٪)	۳۳ (۳۸/۴٪)	۲۵ (۲۹/۱٪)	۱۲ (۱۴٪)			۱. ترقی خواهان	
۱۱ (۱۴٪)	۲۴ (۳۱/۲٪)	۱۶ (۲۰/۸٪)	۲۶ (۳۳/۸٪)				۲. محافظه کاران
۶ (۵/۴٪)	۶ (۵/۴٪)	۴۵ (۴۰/۵٪)	۵۴ (۴۸/۶٪)				۳. زحمت کشان
۲۵ (۳۱/۶٪)	۴۱ (۵۱/۹٪)	۵ (۶/۳٪)	۸ (۱۰/۱٪)	۴. طلایه داران			

خوشه ترقی خواهان بیشترین فراوانی را دارد. نتایج جدول ۷ نشان می‌دهد بیمه‌گذارانی که یک نفر را به منزله استفاده‌کننده از سرمایه در صورت فوت تعیین کرده‌اند در خوشه ترقی خواهان دارای بیشترین فراوانی، بیمه‌گذارانی که دو نفر را تعیین کرده‌اند در خوشه‌های زحمت کشان و طلایه داران و بیمه‌گذارانی که سه نفر و بیشتر را تعیین کرده‌اند در خوشه زحمت کشان بیشترین فراوانی را دارند.

ماهر، بیشترین فراوانی مربوط به خوشه زحمت کشان و در مشاغل کارمندی و صاحبان بنگاه‌ها، بیشترین میزان فراوانی مربوط به اعضای گروه طلایه داران است. نتایج جدول ۶ نشان می‌دهد که در نحوه پرداخت ماهانه، بیشترین فراوانی مربوط به خوشه زحمت کشان، در نحوه پرداخت سه و شش ماهه، خوشه محافظه کاران و در نحوه پرداخت سالانه،

جدول ۶: جدول توافقی متغیر نحوه پرداخت حق بیمه به تفکیک هر خوشه

نحوه پرداخت حق بیمه				
خوشه	ماهانه	سه ماهه	شش ماهه	سالانه
فراوانی (درصد)	۱. ترقی خواهان	۸ (٪۹/۱)	۱۰ (٪۱۱/۶)	۲۶ (٪۳۰/۲)
	۲. محافظه کاران	۲۲ (٪۲۸/۶)	۱۲ (٪۱۵/۶)	۱۰ (٪۱۳)
	۳. زحمت کشان	۱۸ (٪۱۶/۲)	۱۰ (٪۹)	۱۷ (٪۱۵/۳)
	۴. طلایه داران	۱۸ (٪۲۲/۵)	۱۰ (٪۱۲/۷)	۲۲ (٪۲۷/۸)

جدول ۷: جدول توافقی تعداد استفاده کنندگان از سرمایه بیمه در صورت فوت به تفکیک هر خوشه

تعداد استفاده کنندگان سرمایه در صورت فوت				
خوشه	یک نفر	دو نفر	سه نفر و بیشتر	
فراوانی (درصد)	۱. ترقی خواهان	۵۳ (٪۶۱/۶)	۲۷ (٪۳۱/۴)	۶ (٪۷)
	۲. محافظه کاران	۳۶ (٪۴۶/۸)	۲۴ (٪۳۱/۲)	۱۷ (٪۲۲/۱)
	۳. زحمت کشان	۴۲ (٪۳۷/۸)	۳۷ (٪۳۳/۳)	۳۲ (٪۲۸/۸)
	۴. طلایه داران	۲۸ (٪۳۵/۴)	۳۷ (٪۴۶/۸)	۱۴ (٪۱۷/۷)

نتایج جدول ۸ نشان می‌دهد که در نرخ ثابت افزایش سرمایه فوت، بیشترین فراوانی مربوط به خوشه محافظه کاران، در نرخ‌های ۵ درصد و ۱۰ درصد خوشه ترقی خواهان و نرخ‌های ۱۵ درصد و ۲۰ درصد خوشه طلایه داران را دارند. نتایج جدول ۹ نشان می‌دهد که در نرخ ثابت افزایش سرمایه فوت، بیشترین فراوانی مربوط به خوشه زحمت کشان، در نرخ‌های ۵ و ۱۵ درصد خوشه طلایه داران و در نرخ ۲۰ درصد خوشه ترقی خواهان بیشترین فراوانی را دارند.

جدول ۸: جدول توافقی متغیر نرخ افزایش سرمایه فوت به تفکیک در هر چهار خوشه

نرخ افزایش سرمایه فوت					
خوشه	ثابت	۵٪	۱۰٪	۱۵٪	۲۰٪
فراوانی (درصد)	۱. ترقی خواهان	۴۶ (٪۵۳/۵)	۲۶ (٪۳۰/۲)	۱۴ (٪۱۶/۳)	۰ (٪۰)
	۲. محافظه کاران	۶۴ (٪۸۳/۱)	۱ (٪۱/۳)	۱۲ (٪۱۵/۶)	۰ (٪۰)
	۳. زحمت کشان	۱۰۰ (٪۹۰/۱)	۹ (٪۸/۱)	۲ (٪۱/۸)	۰ (٪۰)
	۴. طلایه داران	۵۹ (٪۷۴/۷)	۹ (٪۱۱/۴)	۶ (٪۷/۶)	۳ (٪۳/۸)

جدول ۹: جدول توافقی متغیر نرخ افزایش حق بیمه منظم به تفکیک در هر چهار خوشه

نرخ افزایش منظم حق بیمه					
خوشه	ثابت	۵٪	۱۰٪	۱۵٪	۲۰٪
فراوانی (درصد)	۱. ترقی خواهان	۱۰ (٪۱۱/۶)	۲ (٪۲/۳)	۶۹ (٪۸۰/۲)	۲ (٪۲/۳)
	۲. محافظه کاران	۳۶ (٪۴۶/۸)	۰ (٪۰)	۳۷ (٪۴۸/۱)	۱ (٪۱/۳)
	۳. زحمت کشان	۶۶ (٪۵۹/۵)	۴ (٪۶/۳)	۳۹ (٪۳۵/۱)	۱ (٪۰/۹)
	۴. طلایه داران	۴۱ (٪۵۱/۹)	۵ (٪۶/۳)	۲۶ (٪۳۲/۹)	۴ (٪۵/۱)

۳-۲. ارزیابی کلی خوشه‌ها

بنگاه (۳۱٪) مشغول به کار بوده‌اند و درصد بالایی از آنان حق بیمه خود را به ترتیب به صورت ماهانه (۳۷٪) و سالانه (۲۸٪) پرداخت می‌کنند. اغلب بیمه‌گذاران این خوشه دو نفر (۴۶٪) را برای دریافت سرمایه در صورت فوت انتخاب کرده‌اند. تعداد بسیار بالایی از افراد این خوشه (۷۴٪) نرخ افزایش سرمایه فوت را ثابت انتخاب کرده‌اند. همچنین نزدیک به ۵۰ درصد افراد در این خوشه نرخ افزایش حق بیمه ثابت را انتخاب کرده‌اند.

نتیجه‌گیری

به منظور ارائه چهارچوبی برای بخش‌بندی بیمه‌گذاران بیمه عمر از روش داده‌کاوی و تکنیک الگوریتم کا - میانگین استفاده شد. با توجه به حجم داده‌های در اختیار و نظر متخصص، تعداد خوشه‌ها براساس چهار خوشه آغاز شد. روند اجرا و ارزیابی روش خوشه‌بندی کا - میانگین با توجه به معیار سیلوئت انجام داده شد. نتایج الگوریتم نشان داد که ۲۴ درصد بیمه‌گذاران در خوشه اول، ۲۱ درصد در خوشه دوم، ۳۱ درصد در خوشه سوم و ۲۲ درصد بیمه‌گذاران در خوشه چهارم قرار داشته‌اند و ضریب معیار به دست آمده (۰/۲۸) حاکی از مطلوبیت خوشه‌های به دست آمده است. از آنجاکه متغیرهای ورودی پژوهش شامل ۱۴ متغیر بود، برای تسهیل فرایند بررسی متغیرهای ورودی در هر خوشه، فقط متغیرهای ورودی - که در هر خوشه در آزمون خی دو تفاوت معناداری بین آن‌ها در سطح ۰/۰۱ وجود داشت - وارد تحلیل نهایی شد. به این معنا که در نهایت هفت متغیر سن، تحصیلات، شغل، نحوه پرداخت حق بیمه، تعداد استفاده‌کنندگان از سرمایه بیمه در صورت فوت، نرخ افزایش سرمایه فوت و پرداخت حق بیمه منظم، به منزله متغیرهای ورودی انتخاب و درون هر خوشه در قالب جداول توافقی متغیرهای ورودی خوشه (جدول ۲ تا ۸) ارائه شد.

اگرچه پژوهشی که دقیقاً منطبق با نتایج به دست آمده باشد یافت نشد، اما نتایج پژوهش حاضر با یافته‌های مطالعه باش افشار و همکاران مبنی بر تأثیر متغیرهای جمعیتی (جنس و سن) و بیمه‌ای (نرخ حق بیمه و ضریب فوت) در خوشه‌بندی مشتریان و بهناز و حسینی مبنی بر به دست آمدن چهار گروه پلاتینیومی، طلایی، نقره‌ای و سربی در دسته‌بندی مشتریان به لحاظ سطح استفاده از خدمات شبکه شتاب، شریفی مبنی بر دسته‌بندی میزان سوددهی نوع بیمه برای شرکت بیمه، عبدی و همکاران مبنی بر استفاده از روش کا - میانگین برای تشخیص مشتریان با قابلیت اعتماد بالا و پیش‌بینی میزان فروش پوشش‌های ویژه بیمه و مطالعه فنگ و همکاران مبنی بر پیش‌بینی سودآوری مشتریان بیمه براساس سن، وضعیت بیمه، جنس و نوع شغل مشتری همسو است (باش و همکاران، ۱۳۹۷؛ بهناز و حسینی، ۱۳۹۷؛ شریفی، ۱۳۹۶؛ Abdi et al., 2018; Fang et al., 2017).

۱) خوشه ترقی خواهان: بیشترین درصد فراوانی خوشه اول را بیمه‌گذاران ۳۱ تا ۴۰ سال (۵۹٪) تشکیل داده‌اند که به ترتیب تحصیلات کارشناسی (۳۰٪) و دیپلم (۲۹٪) داشته‌اند و در مشاغل همچون کارمندی (۳۸٪) و کارگر ماهر (۲۹٪) مشغول به کار بوده‌اند و اغلب حق بیمه خود را به ترتیب به صورت ماهانه (۴۹٪) و سالانه (۳۰٪) پرداخت کرده‌اند. اغلب بیمه‌گذاران (۶۱٪) این خوشه، یک نفر را به منزله استفاده‌کننده از سرمایه بیمه در صورت فوت برگزیده و نرخ افزایش سرمایه فوت را ثابت (۵۳٪) انتخاب کرده‌اند. همچنین ۸۰٪ بیمه‌گذاران در این خوشه نرخ افزایش حق بیمه را ۱۰٪ انتخاب کرده‌اند.

۲) خوشه محافظه کاران: بیشترین درصد فراوانی خوشه دوم را بیمه‌گذاران بالای ۵۰ سال (۳۴٪) تشکیل داده‌اند که به ترتیب تحصیلات دیپلم (۳۵٪) و کارشناسی (۳۲٪) داشته‌اند و در مشاغل همچون کارگری ساده (۳۴٪) و کارمند (۳۱٪) مشغول به کار بوده‌اند و اغلب حق بیمه خود را به ترتیب به صورت ماهانه (۴۳٪) و سه ماهه (۲۸٪) پرداخت کرده‌اند. اغلب بیمه‌گذاران (۴۶٪) این خوشه، یک نفر را به منزله استفاده‌کننده از سرمایه بیمه در صورت فوت برگزیده‌اند و نرخ افزایش سرمایه فوت را ثابت (۸۳٪) انتخاب کرده‌اند. همچنین ۴۸ و ۴۶ درصد بیمه‌گذاران در این خوشه، به ترتیب نرخ افزایش حق بیمه ۱۰ درصد و ثابت را انتخاب کرده‌اند.

۳) خوشه زحمت‌کشان: بیشترین درصد فراوانی خوشه سوم را بیمه‌گذاران بین ۴۱ تا ۵۰ سال (۳۴٪) تشکیل داده‌اند که به ترتیب تحصیلات دیپلم (۵۳٪) و زیردیپلم (۴۲٪) داشته‌اند و در مشاغل همچون کارگری ساده (۴۸٪) و کارگری ماهر (۴۰٪) مشغول به کار بوده‌اند و درصد بالایی از آنان حق بیمه خود را به ترتیب به صورت ماهانه (۵۹٪) پرداخت کرده‌اند. تعداد افرادی که بیمه‌گذاران برای دریافت سرمایه در صورت فوت انتخاب کرده‌اند بسیار نزدیک به هم است؛ به گونه‌ای که ۳۷ درصد یک نفر، ۳۳ درصد دو نفر و ۲۸ درصد سه نفر را انتخاب کرده‌اند. تعداد بسیار بالایی از افراد این خوشه (۹۰٪) نرخ افزایش سرمایه فوت را ثابت انتخاب کرده‌اند. همچنین نزدیک به ۶۰ درصد بیمه‌گذاران در این خوشه نرخ افزایش حق بیمه ثابت را انتخاب کرده‌اند.

۴) خوشه طلایه‌داران: نتایج تحلیل توصیفی داده‌ها نشان داد که بیشترین درصد فراوانی خوشه چهارم را بیمه‌گذاران بین ۳۱ تا ۴۰ سال (۳۹٪) تشکیل داده‌اند که به ترتیب تحصیلات کارشناسی (۵۵٪) و تحصیلات تکمیلی (۲۱٪) داشته‌اند و به ترتیب در مشاغل دولتی و خصوصی (کارمند ۵۱٪) و صاحبان

سیاست‌گذاری واحد به علت تفاوت در ویژگی‌های متفاوت بیمه‌گذاران وجود ندارد، بنابراین لازم است که متناسب با هر گروه، سیاست مناسب با همان گروه ارائه شود. در پژوهش حاضر، طی بررسی‌های به‌عمل‌آمده، بیمه‌گذاران بیمه عمر به چهار خوشه تقسیم شدند که هر خوشه ویژگی منحصر به فرد خود را داشتند.

۱. خوشه ترقی‌خواه: با توجه اینکه افراد حاضر در این خوشه جزو جمعیت جوان با تحصیلات بالا هستند، با روش‌های نوین بازاریابی، مانند بازاریابی شبکه‌ای و ارائه وام با نرخ بهره مناسب و بدون ضامن و وثیقه و دریافت سود مشارکت در منافع به میزان ۸۵ درصد سود مازاد محقق شده در سرمایه‌گذاری شرکت که در آیین‌نامه ۶۸ بیمه مرکزی (۱۳۹۰) نیز ارائه شده است، می‌توان بیمه‌گذار را تشویق به سرمایه‌گذاری در بیمه عمر کرد؛

۲. خوشه محافظه‌کاران: با توجه اینکه اغلب بیمه‌گذاران حاضر در این خوشه از جمله بیمه‌گذاران پا به سن گذاشته و تحصیل کرده‌ای هستند که وضعیت اقتصادی متوسطی دارند می‌توان با ارائه پوشش درمانی امراض و بیماری‌های خاص، کمک هزینه معینی را برای درمان بیماری‌های بیمه‌گذاران در نظر گرفت. همچنین معافیت از پرداخت حق بیمه پوشش‌های بیمه‌ای در صورت از کارافتادگی کامل (به دلایلی اعم از حادثه یا بیماری) در جذب و تشویق این دسته از بیمه‌گذاران مؤثر است؛

۳. خوشه زحمت‌کشان: اغلب بیمه‌گذاران حاضر در این خوشه، از جمله بیمه‌گذاران میان‌سالی هستند که میزان تحصیلات و سطح سواد نسبتاً پایینی دارند و اغلب در مشاغل کارگری مشغول به کارند؛ بنابراین به‌کارگیری سیاست‌هایی همچون انعطاف‌پذیری در پرداخت حق بیمه و تعیین سرمایه بیمه‌ای، پرداخت سود تشویقی و پرداخت سود روزشمار، دریافت تسهیلات ویژه در قالب وام بدون بهره با اقساط بلندمدت و امکان دریافت سه وام یا بیشتر (در بیمه‌های عمر فقط امکان دریافت دو وام در طول مدت قرارداد وجود دارد)، می‌تواند در جذب این دسته از بیمه‌گذاران مؤثر باشد.

۴. خوشه طلایه‌داران: اغلب افراد حاضر در این خوشه، افراد با تحصیلات بالا هستند و مشاغل با درآمد مناسبی دارند که می‌توان با ارائه سیاست‌هایی همچون تشکیل اندوخته سرمایه‌گذاری از محل حق بیمه با سود تضمینی، رفع اثر منفی تورم و معافیت از مالیات بر درآمد و وارث (از محل سرمایه و مبالغ پرداختی بیمه به فرد) و ارائه خدمات بیمه به‌صورت الکترونیکی (در قالب فروش اینترنتی، بهره‌مندی از تخفیفات اینترنتی و یا بازخریدی اینترنتی) سعی در جذب سرمایه آنان کرد.

از جمله محدودیت‌های پژوهش حاضر، تعمیم‌یافتن نتایج به دوره‌های گوناگون زمانی (تغییر شرایط روند نیاز بازار، شرایط اقتصادی مانند نرخ تورم، تغییر سیاست‌های کلان در صنعت بیمه) و مکانی (مشتریان بیمه در شهر اصفهان) است. همچنین استفاده

در تبیین احتمالی نتایج خوشه ۱ و علت نام‌گذاری آن تحت عنوان ترقی‌خواهان، گفتنی است که مشتریان حاضر در این خوشه، اغلب جوانان تحصیل‌کرده و سرپرستان خانواده‌اند که خواهان سرمایه‌گذاری مطمئن و سودآوری هستند که بتوانند آینده خود و خانواده‌شان را تضمین کنند؛ بنابراین با ارائه سود تضمینی و مشارکت در منافع به این گروه مشتریان و ارائه آگاهی از سودآوری سرمایه‌گذاری صورت‌گرفته برای سرپرست خانواده و اعضا، می‌توان بازار نسبتاً مناسبی برای این گروه سنی خلق کرد.

در تبیین نتایج خوشه ۲ و علت نام‌گذاری آن تحت عنوان محافظه‌کاران، گفتنی است که سن یکی از عوامل مؤثر در قدرت ریسک‌پذیری است. با افزایش سن، قدرت ریسک‌پذیری در افراد کاهش می‌یابد و تمایل به بازگشت اندوخته (برای تأمین هزینه‌های درمان یا کمک به خرید مسکن برای فرزندان یا ازدواج آنان) افزایش می‌یابد. شرکت بیمه می‌تواند با پیشنهاد دادن بندهای مکمل، مانند استفاده از خدمات درمانی و پوشش‌های درمانی خاص در قالب پوشش اضافی بکوشد تا نظر این دسته از بیمه‌گذاران را به خود جلب کند.

در تبیین نتایج خوشه ۳ و علت نام‌گذاری آن تحت عنوان زحمت‌کشان گفتنی است که اغلب افراد حاضر در این خوشه در مشاغل کارگری مشغول به کارند و به طبع آن به نسبت سایر خوشه‌ها سطح درآمد نسبتاً پایین‌تری دارند. با توجه به سرمایه اندک این گروه، می‌توان شرایط را به گونه‌ای فراهم کرد که این دسته از بیمه‌گذاران نیز بتوانند با رعایت حداقل‌های تعیین‌شده میزان حق بیمه و سرمایه بیمه‌ای مدنظر خود را تعیین کنند.

همچنین در تبیین نتایج خوشه ۴ و علت نام‌گذاری آن تحت عنوان طلایه‌داران شایان ذکر است که میزان درآمد بالا و به طبع آن، میزان سرمایه بالایی که این قشر را قادر به سرمایه‌گذاری می‌کند از جمله فاکتورهای جذابی است که مدیران شرکت‌های بیمه را به ارجحیت این خوشه به سایر خوشه‌ها متمایل کند. همچنین سطح تحصیلات بالا نیز یکی دیگر از ویژگی‌های مثبت این گروه است. به باور دامغانیان، افراد با تحصیلات بالاتر به نسبت افراد با تحصیلات پایین‌تر، از بیمه عمر استقبال بیشتری می‌کنند. در حقیقت تولیدکنندگان، صاحبان بنگاه‌ها و شاغلان با تحصیلات بالا، آگاهی بالاتری درباره کارکردها و خدمات بیمه دارند و این دسته از بیمه‌گذاران می‌توانند در اولویت‌های بازاریابی قرار گیرند (دامغانیان، ۱۳۸۶).

پیشنهاد‌های سیاستی

راهبرد کلی شرکت بیمه پارسیمان توسعه آرامش و امنیت خاطر برای جامعه با ارائه خدمات بیمه‌ای و شرکت در بازارهای سرمایه‌ای و مالی به‌منظور تأمین سودآوری مطمئن است که این مهم با فروش بیشتر و در نتیجه سود بیشتر حاصل می‌شود و سیاست‌گذاری شرکت نیز بر همین اساس است. با توجه به اینکه امکان ارائه

تعاملی اطلاعات، کیش، دانشگاه تهران.

شریفی، سمیه (۱۳۹۶). «دسته‌بندی مشتریان بیمه براساس داده‌کاوی بیمه جهت تعیین سودآورترین نوع بیمه (مطالعه موردی: بیمه سامان)». پایان‌نامه کارشناسی ارشد رشته مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد صفهان.

شیرخدايي، میثم، نجات، سهیل و اکبری، امیر (۱۳۹۵). «بخش‌بندی بازار بیمه عمر از طریق سبک زندگی مشتریان با استفاده از الگوی AIO». تحقیقات بازاریابی نوین، دوره ششم، شماره ۲۱، ص ۷۴-۵۵.

طاهری، آزاده (۱۳۹۷). «بیمه مسئولیت مدنی در فضای مجازی». پایان‌نامه کارشناسی ارشد، رشته حقوق خصوصی، دانشگاه آزاد اسلامی خوراسگان.

عابدی، مرتضی (۱۳۹۵). «رابطه بین فرهنگ طبقاتی مردم با خرید انواع بیمه در شهر اصفهان». پایان‌نامه کارشناسی ارشد دانشگاه آزاد واحد خوراسگان.

عاملی بصیری، مرضیه و قره‌خانی، محسن (۱۳۹۴). «طراحی و پیاده‌سازی سیستم پایلوت فروش مقاطع با استفاده از رویکرد ارزش مشتری در صنعت بیمه ایران». فصلنامه پژوهشنامه بیمه، سال سی‌ام، شماره ۳ پیاپی ۱۱۹، ص ۲۱-۴۰.

عباسی، زینب و محسن، حسنی (۱۳۹۷). «شناسایی تأثیر ریسک‌های اقتصادی بر تقاضای بیمه عمر و عملکرد مالی شرکت‌های بیمه». نخستین همایش ملی حسابداری و مدیریت، نطنز، دانشگاه آزاد اسلامی واحد نطنز.

قره‌خانی، محسن و ابوالقاسمی، مریم (۱۳۹۴). «کاربردهای داده‌کاوی در صنعت بیمه». تازه‌های جهان بیمه، سال چهاردهم، شماره ۱۵۸ و ۱۵۹، ص ۵-۲۲.

کریمی، سیدمحمد (۱۳۹۰). «ارزیابی عملکرد صنعت بیمه کشور و تبیین چشم‌انداز آینده». فصلنامه سیاست‌های مالی و اقتصادی، سال اول، شماره ۲، ص ۱۸۳-۲۰۲.

مؤمنی، منصور (۱۳۹۰). خوشه‌بندی داده‌ها (تحلیل خوشه‌ای). تهران: انتشارات مؤلف.

Abdi, F., Khalili-Damghani, K. and Abolmakarem, S. (2018). "Solving customer insurance coverage sales plan problem using a multi-stage data mining approach". *Kybernetes*, 47(1), pp. 2-19.

Amorim, R. C. and Komisarczuk, P. (2012). On initializations for the Minkowski Weighted K-Means. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7619 LNCS, 45-55.

Bingham, J. R., Umphress, E.E. and Martin, C. (2005). *Business Marketing*. New York: McGrawHill.

Blonder, N., Orsburn, BC., Blonder, J. and Gonzalez, C. A. (2019). "Visual mass-spec share (vMS-Share): a new public web-based mass spectrometry visualization and data mining repository". *J Proteomics Bioinform*, 12(2), pp. 6-15.

از دیگر روش‌های خوشه‌بندی می‌تواند نتایج را دستخوش تغییراتی کند که در این مسیر، به پژوهشگران علاقه‌مند پیشنهاد می‌شود به‌منظور خوشه‌بندی و تجزیه و تحلیل داده‌های بیمه‌گذاران، از روش‌هایی نظیر روش کوهن، شبکه‌های عصبی، خوشه‌بندی سلسله‌مراتبی و TwoStep استفاده کنند که با ارزیابی و مقایسه نتایج این پژوهش‌ها با یکدیگر، می‌توان به غنی‌کردن بیشتر ادبیات این حوزه کمک کرد.

منابع

آذر، عادل، احمدی، پرویز و سبط، محمدوحید (۱۳۸۹). «طراحی مدل انتخاب نیروی انسانی با رویکرد داده‌کاوی (مورد مطالعه: استخدام داوطلبان آزمون‌های ورودی یک بانک تجاری در ایران)». مدیریت فناوری اطلاعات، سال دوم، شماره ۳، ص ۲۲-۳.

اکبری، امیر، شیرخدايي، میثم و نجات، سهیل (۱۳۹۵). «بخش‌بندی بازار بیمه عمر از طریق سبک زندگی مشتریان با استفاده از الگوی AOI». تحقیقات بازاریابی نوین، سال ششم، شماره ۲۱، ص ۷۴-۵۵.

باش افشار، مریم، سعیدپناه، مسعود و تیره عیدوزهی، فرشید (۱۳۹۷). «الگوی خوشه‌بندی مشتریان بیمه عمر (مطالعه موردی: یک شرکت بیمه‌ای)». پژوهشنامه بیمه، دوره ۳۳، شماره ۲، ص ۴۵-۶۴.

بهروزیان، محمد (۱۳۹۵). «دسته‌بندی خودکار متون با استفاده از تکنیک‌های داده‌کاوی». دومین همایش ملی کامپیوتر.

بهناز، شهرزاد و حسینی، راحیل (۱۳۹۷). «دسته‌بندی سرویس‌های مشتریان به لحاظ سطح استفاده از خدمات شبکه‌شتاب با استفاده از مدل‌های ترکیبی». پژوهش‌های نوین در تصمیم‌گیری، سال سوم، شماره ۴، ص ۷۰-۵۱.

بیمه مرکزی ایران (۱۳۹۰). آیین‌نامه بیمه‌های زندگی و مستمری. تهران: انتشارات بیمه مرکزی ایران.

ترکی هرچگانی، نادر (۱۳۹۴). «راهکارهای گسترش بیمه‌های عمر و پس‌انداز در ایران از منظر نمایندگان شرکت سهامی بیمه ایران شهر اصفهان». پایان‌نامه کارشناسی ارشد مدیریت دولتی، دانشگاه آزاد واحد خوراسگان.

حلاج‌نژادی، سیدهادی (۱۳۹۷). «موضوع بیمه، مطالعه موردی: نگاهی عمیق‌تر به بیمه عمر». چهارمین سمپوزیوم بین‌المللی علوم مدیریت، تهران، مؤسسه آموزشی عالی مهر اروند و مرکز راهکارهای دستیابی به توسعه پایدار.

دامغانیان، حسین (۱۳۸۶). «بازاریابی/ بیمه عمر (انفرادی)». پژوهشنامه بیمه، شماره ۴۵، ص ۱۰۶-۱۰۹.

رجوعی، مرتضی و شیعه‌زاده، الهه (۱۳۸۷). «تعیین معیارهای جذابیت بازار و انتخاب بازار هدف». فصلنامه علمی تخصصی دانشگاه امام رضا، شماره ۵، ص ۱۲-۵.

شادکام اصفهانی، شهرام (۱۳۹۵). «بررسی عوامل مؤثر بر عدم توسعه فرهنگ بیمه در شهرستان اصفهان». پایان‌نامه کارشناسی ارشد مدیریت بیمه دانشگاه آزاد اسلامی خوراسگان.

شاهینی، شبنم، تربتی، مهشید و ایزدی، عبدالرضا (۱۳۹۵). «داده‌کاوی در پایگاه داده‌های بزرگ: مفاهیم و روش‌ها». اولین کنفرانس بین‌المللی بازاریابی

- Boodhun, N. and Jayabalan, M. (2018). "Risk prediction in life insurance industry using supervised learning algorithms". *Complex & Intelligent Systems*, 4(2), pp. 145-154.
- Camilleri, M. A. (2017). *Market Segmentation, Targeting and Positioning BT - Travel Marketing, Tourism Economics and the Airline Product: An Introduction to Theory and Practice*. Springer International Publishing.
- Celebi, E. M., Kingravi, H. and Vela, P. A. (2013). "A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm". *Expert Systems with Applications*. 40(1), pp. 200-210.
- Chien-Wen, H. (2012). "Using the Taguchi method for effective market segmentation". *Expert Systems with Applications*, 39(5), pp. 5451-5459.
- Destpand, S. and Thaker, I. (2010). "Appropriate medical data categorization for data mining classification techniques". *Medical Informatics and the Internet in Medicine*, 27(1), pp. 59-67.
- Fang, K., Jiang, Y. and Song, M. (2016). "Customer profitability forecasting using Big Data analytics: A case study of the insurance industry". *Computers & Industrial Engineering*. 101, pp. 554-564.
- Goetzke, B. I. and Spiller, A. (2014). "Health-improving lifestyles of organic and functional food consumers". *British Food Journal*, 3(116), pp. 4-12.
- Hand, D. J., Smyth, P. and Mannila, H. (2011). *Principles of Data Mining*. Cambridge, MA, USA: MIT Press.
- Kaigorodova, G. N. and Mustafina, A. A. (2014). "The influence of forms of insurance coverage organization on population's life quality". *Mediterranean Journal of Social Sciences*, 5(24), pp. 118-123.
- Lamberti, L. and Noci, G. (2009). "Marketing strategy and marketing performance measurement system: Exploring the relationship". *European Management Journal*, 28(2), pp. 139-152.
- Mahalakshmi, V. and Chitra, D. (2014). "An Empirical Study on Lifestyle Segmentation on the Retail Store Attributes that are Important to Old People". *Asia Pacific Journal of Marketing & Management Review*, 19(2), pp. 92-105.
- Marbin, M., Seifuddin, F., Judy, J., Mahon, P. B., Bipolar Genome Study (BiGS) Consortium, Potash, J. B. and Zandi, P. P. (2012). "Data mining approaches for genome-wide association of mood disorders". *Psychiatric genetics*, 22(2), pp. 55-61.
- McDonald, M. and Dunbar, I. (2012). *Market Segmentation: How to do it and how to profit from it*, Revised 4 th Edition. John Wiley and Sons.
- Trindade, G., Dias, J. and Ambrosio, J. (2017). "Extracting clusters from aggregate panel data: A market segmentation study". *Applied Mathematics and Computation*, 296, pp. 277-288.
- Wu, Z., Lin, w., Zhang, Z., Wen, A., and Lin, L. (2017). "An Ensemble Random Forest Algorithm for Insurance Big Data Analysis". International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, 5(1), pp. 531-536.

Presenting a Framework for Segmentation of Life Insurance Customers Using Data Mining

Maryam Roostazadeh Sheikh Yousefi
Seid Mohammad Reza Mirahmadi

Abstract

The purpose of this study is to provide a framework for segmentation of life insurance customers using data mining techniques. The statistical population of the study consisted of customers of an insurance company in Isfahan, where the required data were collected from contracts of life insurance during the years 2008 to 2018. Data were collected on 353 life insurance policyholders in 14 variables (in terms of individual characteristics and selected insurance conditions). Customers were classified into 4 clusters using K-means clustering algorithm and Matlab software. The analysis of the results of the customer behavior of each cluster provided a basis for naming the clusters as progressives, conservatives, toilers, and vanguards. Also, to facilitate the process of examining the input variables in each cluster, only 7 variables (age, education, occupation, and premium payment, number of beneficiaries of death insurance, death rate increase and regular premium payment) are included in each cluster. In the Chi-square test, there were two significant differences at the 0.001 level.

Keywords: Customer Segmentation, Data Mining, K-means Algorithm, Life Insurance