

مدیریت دانش مشتریان بانک با استفاده از روش خوشه‌بندی Anti-Harmonic-Means

سیده عطیه علی‌کیا امیری^{*}، جواد وحیدی^۲

۱- کارشناس ارشد مهندسی فناوری اطلاعات، ۲- استادیار، دانشگاه علم و صنعت ایران

(دریافت: ۱۳۹۷/۰۸/۲۵ پذیرش: ۱۳۹۸/۰۷/۰۲)

چکیده

افزایش روزافزون سطح رقابت در بازار، مدیران و تحلیلگران سازمان‌ها را وادار ساخته به دنبال راه‌کارهایی باشند که مزیت رقابتی را برای سازمان به ارمغان آورند. بر اساس مطالعات موجود، استفاده از دانش مشتری جهت اتخاذ راهبردهای لازم برای جلب رضایت مشتری می‌واند سازمان‌ها را به سمت تحقق این هدف سوق دهد. از سوی دیگر گسترش بهره‌گیری از فناوری‌های بروز در زمینه اطلاعات و ارتباطات خصوصاً در بانک‌ها باعث بر جای ماندن حجم عظیمی از داده‌ها گشته که تحلیل و تصمیم‌گیری بر اساس آن‌ها با روش‌های معمول گزارش‌گیری و روش‌های آماری امکان‌پذیر نمی‌باشد. داده‌کاوی ابزار بروز و قدرتمندی است که در این تحقیق جهت تحلیل داده‌ها به جهت استخراج دانش مشتری پیشنهاد می‌گردد.

هدف از این تحقیق بخش‌بندی مشتریان بانک ملت، باهدف کشف ویژگی‌های رفتاری مشابه، برای کمک به مدیران این بانک جهت تسهیل اتخاذ استراتژی‌های متناسب با هر بخش و در نتیجه حفظ، تقویت و یا توسعه ارتباط با مشتریان و نهایتاً سودآوری برای این بانک می‌باشد. در این مقاله از روش Anti Harmonic Means برای خوشه‌بندی مشتریان استفاده شده و داده‌های خام موردنیاز جهت این مطالعه از پایگاه داده‌های بانک ملت استخراج شده است.

واژه‌های کلیدی: مدیریت دانش مشتری، کشف دانش، داده‌کاوی، خوشه‌بندی مشتریان، RFM، A-H-Means

۱- مقدمه

مشتریان، در نگاه اول بسیار سخت و گاهی ناممکن به نظر برسد، اما با بهره‌گیری از ابزارهای نوین تحلیل داده‌ها همچون داده‌کاوی و استفاده از دانش‌های روز دنیا همچون مباحث مدیریت دانش و مدیریت ارتباط با مشتری می‌توان این داده‌ها را به گنجینه‌ای گران‌بها تبدیل نمود. این گنجینه همان دانش استخراج‌شده از طریق کندوکاو این داده‌هاست.

فرایند داده‌کاوی به سازمان‌ها کمک می‌کند تا انبوه داده‌های مشتریان را تحلیل نمایند و اطلاعات مفید را از آن‌ها استخراج نموده و مزیت رقابتی را نسبت به دیگران کسب نمایند [۵]. می‌توان روش‌های داده‌کاوی را ابزارهای پشتیبان تصمیم دانست که مدیران سازمان را قادر می‌سازد که به‌جای تأثیرپذیری از محیط، بر بازار و عوامل آن تأثیرگذار باشند.

رتبه‌بندی مشتریان با بخش‌بندی آن‌ها سعی در کشف الگوهای رفتاری مشابه حاکم بر آن بخش دارد. کشف این الگوها می‌تواند در مسیر اتخاذ تصمیمات راهبردی سازمان بکار گرفته شود.

مدیریت دانش مشتری با ترکیب هر دو رویکرد فناوری محور و داده‌گرا در بحث مدیریت ارتباط با مشتری و رویکرد فردگرا در بحث مدیریت دانش با هدف بهره‌گیری از پتانسیل هم‌افزایی هر دوی آن‌هاست [۳]. نتیجه این مسئله می‌تواند توصیف شیواتر دانش برای مشتری، دانش درباره مشتری و دانش از سوی مشتری باشد. بنابراین، این امکان وجود خواهد داشت که محصولات و خدمات جدید به گروه مناسب از مشتریان تحویل شود. بنابراین ریسک شکست محصول جدید کاهش خواهد یافت [۱۳]. داده‌های مشتریان و ابزارهای فناوری اطلاعات زمینه‌ای را برای ایجاد ارتباط موفق با مشتری ایجاد می‌کنند. علاوه بر این، سرعت روبه‌رشد اینترنت و فناوری‌های وابسته باعث افزایش فرصت بازاریابی شده است و راه‌های ارتباط بین سازمان و مشتریان را تغییر داده است [۴].

هرچند ممکن است استخراج دانش از حجم عظیم داده‌های برجای‌مانده از تعاملات سازمان‌ها و به‌خصوص تعاملات بانک‌ها با

۲- روش تحقیق

۲-۱- کارهای مرتبط

همسیه در [۶] مدلی ادغامی از داده‌کاوی و مدل رتبه‌بندی مشتریان جهت مدیریت کارت‌های اعتباری مشتریان در بانک ارائه می‌دهد. در این تحقیق جهت ساخت مدلی برای شناسایی الگوی رفتاری مشتریان از شبکه‌های عصبی و روش خود سازمان دهنده، استفاده شده است.

چن و همکاران در [۷] به دسته‌بندی مشتریان جهت کشف تقلب و سوءاستفاده‌های مالی و شناسایی بعضی از عوامل کلیدی برای تشخیص تقلب خریداران و فروشندگان در معاملات آنلاین توسط کارت اعتباری پرداخته‌اند. به این منظور از رگرسیون لاجیت و داده‌های مربوط به یک دوره سه ماهه مزایدهات آنلاین فروش لپ‌تاپ استفاده شده است.

مقاله [۸] به خوشه‌بندی مشتریان با خصوصیات رفتاری مشابه (RFM) با هدف تخمین وفاداری مشتری جهت تسهیل اتخاذ استراتژی و پیشنهاد سرویس‌های متناسب هر گروه با ترکیب روش‌های داده‌کاوی و ارزش طول عمر مشتری (CLV)، پرداخته است. در این تحقیق از داده‌های جمعیت شناختی و داده‌های مربوط به تراکنش‌های مالی مشتریان (RFM) مربوط به شرکت ایرانی ساپکو به‌عنوان ورودی و ترکیب روش RFM وزن‌دار با الگوریتم K-Means استفاده شد.

مؤمنی و همکاران در [۱۱] به لزوم ارتباط با مشتری در بانک‌ها پرداخته‌اند. در این مقاله اشاره شده که طبق قاعده معروف ۲۰/۸۰ پارتو معمولاً ۲۰٪ مشتریان، ۸۰٪ سودآوری سازمان را منجر می‌شوند. بنابراین، دسته‌بندی مشتریان می‌تواند ابزار شناسایی این گروه طلایی و ارائه خدمات بیشتر به آن‌ها بوده و سود بیشتر سازمان را برای سازمان به ارمغان آورد.

حسینی در [۱۴] از روش‌های هوشمند داده‌کاوی جهت مدیریت ارتباط با مشتریان بانک پارسیان بهره گرفته است. او در پایان‌نامه خود با ترکیب دو دسته از متغیرها، ۳۰ متغیر برای بررسی الگوی رفتاری مشتریان شناسایی نموده و با استفاده از شبکه عصبی کوهنن ماتریس نرون‌های خروجی شبکه ۳*۵ را تشخیص داده و مشتریان را در ۱۳ خوشه تقسیم نمود. کشف این مسئله که کدام دسته از مشتریان سودمندی لازم را برای نگهداری دارند و جلوگیری از هزینه‌های اضافی تبلیغات به‌صورت گسترده با ارائه سرویس‌های مناسب به مشتریان بالقوه، از نتایج این بررسی اعلام شده است.

حسینی و همکاران در [۱] به بانکداری به‌عنوان موضوعی که دائماً

تحت تأثیر داده‌های بزرگ بوده اشاره می‌کنند و معتقدند بررسی‌های کاربردی موجود تنها برنامه کاربردی تا سال ۲۰۱۳ را پوشش می‌دهند. این مطالعه با هدف جمع‌آوری این شکاف تحقیقی و نمایش پیشرفت‌های قابل توجه و جدیدترین پیاده‌سازی داده‌کاوی در بانکداری بعد از سال ۲۰۱۳ انجام شده است. با جمع‌آوری و تجزیه و تحلیل روند تمرکز تحقیق، منابع داده‌ها، ابزار تحلیلی داده‌ها، این مقاله بینش ارزشمندی در رابطه با پیشرفت‌های آینده داده‌کاوی و بخش بانکی همراه با یک جدول مرجع یکپارچه جامع به دست می‌دهد. علاوه بر این، در این تحقیق موانع کلیدی شناسایی شده و خلاصه‌ای را برای همه افراد علاقه‌مند مواجه با چالش‌های بزرگ داده‌ها ارائه می‌دهد.

دانگ و همکاران در [۹] نقش مدیریت ارتباط با مشتری و داده‌کاوی و توسعه آن‌ها را بررسی می‌کنند. در این مطالعه اشاره شده که با تحلیل‌های حاصل از الگوریتم K-Means بر روی خلاصه اطلاعات تعداد زیادی از مشتریان، می‌توان اطلاعات بسیار مفیدی از آن‌ها را در دسترس مشتریان قرار داد. هدف دسته‌بندی کمک به بانک‌های تجاری برای اجرای CRM که خود اهمیت ویژه‌ای دارد بیان شده است.

در مقاله [۱۰] به داده‌کاوی به‌عنوان ابزاری مؤثر در سیستم CRM بانک‌های تجاری اشاره شده که اطلاعات را به دانش تبدیل می‌کند. داده‌کاوی با کمک به تصمیم‌گیری بهتر برای بانک مفید خواهد بود. در این مقاله تحلیل ویژگی‌های خوشه‌های مشتری، تحقیق روی مدل خرید و کشف تقلب از مهم‌ترین کاربردهای داده‌کاوی بیان شده است.

با توجه به کاربرد مختلف ابزار داده‌کاوی مرتاضی و پارسا در مقاله [۲] از ابزار داده‌کاوی جهت تشخیص شبکه بات و تحلیل رفتاری جریان شبکه بهره گرفتند. بعد از تهیه مجموعه داده‌های واقعی که ترکیبی از ترافیک سالم و ترافیک تولیدی توسط چندین شبکه بات مشهور بود، جریان بسته‌ها در پنجره‌های زمانی ۳۰۰ ثانیه‌ای تحلیل شده و با توجه به الگوهای ترافیکی قابل تشخیص، خصیصه‌های مختلفی استخراج (تولید) شد. آن‌ها این خصیصه‌ها در ابزار وکا و به کمک الگوریتم‌های یادگیری ماشین داده‌کاوی نموده و نتایج طبقه‌بندی را به‌عنوان خروجی ارائه کردند.

۲-۲- الگوریتم Anti Harmonic Means

الگوریتم خوشه‌بندی K-Means یک روش خوشه‌بندی ساده و سریع است که به دلیل پیاده‌سازی آسان و تعداد تکرار کم، عموماً مورد استفاده قرار می‌گیرد. الگوریتم K-Means در تلاش برای یافتن مراکز خوشه‌های (C_1, C_2, \dots, C_K) به‌گونه‌ای عمل می‌کند که مجموع مربعات فاصله‌ی هر نقطه X_i تا نزدیک‌ترین مرکز

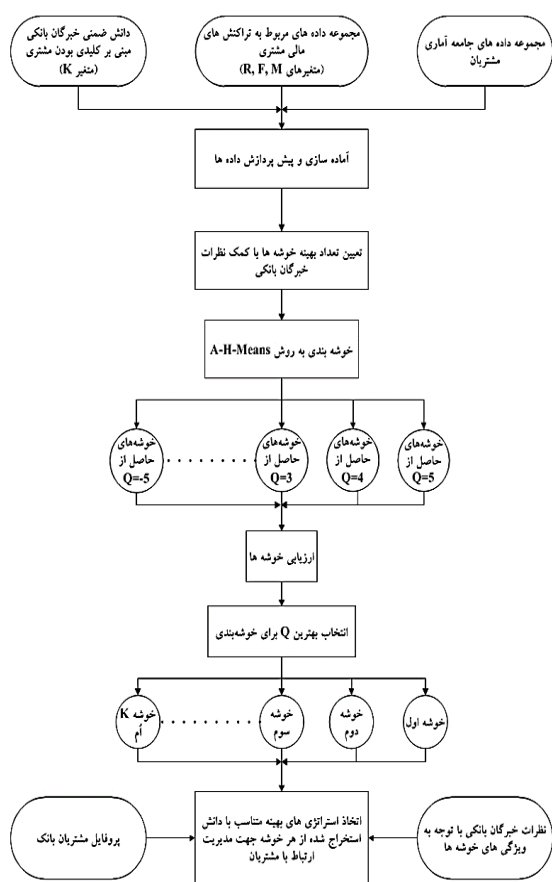
محاسبه می‌شود:

$$c_j = \frac{\sum_{i=1}^n m(c_j|x_i)w(x_i)x_i}{\sum_{i=1}^n m(c_j|x_i)w(x_i)} \quad (4)$$

۵- گام‌های ۳ و ۴ را تا همگرا شدن دسته‌ها تکرار نمایید [۸].

۲-۳- چارچوب تحقیق

در شکل (۱) چارچوب کلی این تحقیق نشان داده شده است.



شکل (۱): چارچوب تحقیق

در ادامه گام‌های مربوط به چارچوب این تحقیق، در مطالعه موردی بانک ملت به کار گرفته خواهد شد.

۲-۳-۱- گام انتخاب و جمع‌آوری متغیرهای ورودی

در حوزه بخش‌بندی رفتاری مشتریان بانک و بررسی رویکردهای سازمانی بانک ملت، سه دسته متغیر داده‌های جمعیت‌شناختی مشتریان شامل جنسیت، سن، تحصیلات، وضعیت تأهل، شغل و تاریخ تولد، محل منطقه بانکی و ... داده‌های مربوط به تراکنش‌های مالی مشتریان نظیر تراکنش اخیر مشتری (R)، متغیر پولی (F)، فراوانی (M) و فاکتور (K) که نشان‌دهنده

خوشه (C_j) کمترین شود. وابستگی کارایی K-Means روی مقاردهای اولیه مراکز، یک مشکل اصلی این الگوریتم می‌باشد. در این الگوریتم ارتباطی قوی بین نقاط داده و نزدیک‌ترین مراکز خوشه برقرار شده و باعث می‌شود مراکز خوشه‌ها از محدوده تراکم محلی داده‌ها خارج نشوند. روش K-harmonic means این مشکل عمده را از طریق جایگزینی کمترین فاصله یک نقطه از مراکز که در K-Means استفاده می‌شود با میانگین هارمونیک فاصله هر نقطه تا تمامی مراکز برطرف می‌کند. میانگین هارمونیک یک امتیاز مناسبی را به هر نقطه‌ای داده بر اساس نزدیکی آن به هر مرکز می‌دهد که این امر را به‌عنوان یک ویژگی میانگین هارمونیک در نظر می‌گیرند. روش A-H-Means، یک روش میانگین‌گیری بسط یافته است که به ازای مقادیر مختلف Q سایر روش‌های محاسبه میانگین را نیز نتیجه خواهد داد.

نمادهای زیر برای فرمول‌بندی الگوریتم A-H-Means استفاده می‌شود:

$$X = \{x_1, x_2, \dots, x_n\}$$

$$C = \{c_1, c_2, \dots, c_k\}$$

Q: متغیر توان که می‌تواند از بازه R انتخاب می‌شود.

۱- الگوریتم با انتخاب K مرکز آغاز می‌شود. این مراکز می‌تواند به‌طور حدسی یا تصادفی انتخاب گردد.

۲- محاسبه مقدار تابع هدف به‌صورت زیر است:

$$AHM(X, C) = \frac{\sum_{i=1}^n \|x_i - c_j\|^{Q+1}}{\sum_{i=1}^n \|x_i - c_j\|^Q} \quad (1)$$

۳- برای هر داده x_i، عضویت m(c_j|x_i) به ازای هر مرکز C_j و وزن مربوط به آن (w(x_i)) محاسبه می‌شود.

تابع عضویت به‌صورت زیر محاسبه می‌شود:

$$m_{AHM}(c_l|x_i) = \begin{cases} 1 & ; \text{if } l = \text{argmin}_j \|x_i - c_j\|^2 \\ 0 & ; \text{otherwise} \end{cases} \quad (2)$$

تابع وزن عضویت به‌صورت زیر محاسبه می‌شود:

$$w_{AHM}(x_i) = 1 \quad (3)$$

۴- برای هر دسته، مجدداً مرکز C_j جدید به‌صورت زیر

است.

۲-۳-۲- گام آماده‌سازی و پیش‌پردازش داده‌ها

داده‌های این تحقیق از داده‌های موجود در پایگاه اطلاعاتی بانک ملت جمع‌آوری شده است. از آنجایی که داده‌های خام انتخاب مناسبی برای اعمال الگوریتم‌ها نیستند نیاز است عملیاتی بر روی داده‌ها جهت آماده‌سازی آن‌ها برای ورود به الگوریتم‌ها صورت گیرد. از جمله این عملیات نرمال‌سازی داده‌ها می‌باشد.

از میان یازده متغیر شناسایی شده جهت مطالعه رفتار مشتریان، متغیرهای تراکنش اخیر مشتری و متغیر کلیدی بودن مشتری به‌عنوان ورودی الگوریتم‌های خوشه‌بندی در نظر گرفته می‌شوند. اما از آنجایی که ویژگی‌های انتخابی از نمونه‌های مورد مطالعه در این تحقیق دارای بازه تغییرات یکسان نمی‌باشند، برای جلوگیری از تأثیر متفاوت ویژگی‌ها (از آنجایی که متغیرهای با مقادیر بزرگ تأثیر بیشتری بر تابع هدف می‌گذارند که لزوماً به معنی مهم‌تر بودن آن‌ها در الگوریتم بخش‌بندی نیست) از نرمال‌سازی به روش Min-Max (فرمول ۵) استفاده شده است.

$$X^* = \frac{X - \text{Min}(X)}{\text{Range}(X)} \quad (5)$$

در فرمول بالا X نشان‌دهنده مقدار متغیر است.

به‌طور مثال جدول (۲) یک نمونه ده‌تایی از متغیرهای نرمال شده R، F، M و K از داده‌های مورد مطالعه در این تحقیق را نشان می‌دهد.

جدول (۲): بخشی از متغیرهای نرمال شده

ID	R	F	M	K
۴۱۱	۰/۰۳۵۲۹۴	۰	۰/۳۹۴۷۲۲۸۵۲	۰
۴۱۲	۰/۸	۰/۶۸۹۶۵۵۱۷۲	۰/۴۶۶۰۱۰۷۱	۰
۴۱۳	۰/۲۰۵۸۸۲	۰	۰/۷۷۰۶۱۶۲۹۵	۰
۴۱۴	۰/۱	۰/۶۲۰۶۸۹۶۵۵	۰/۹۰۱۹۳۰۱۶	۰
۴۱۵	۰/۰۷۰۵۸۸	۰	۰/۲۱۳۲۶۵۹۱۱	۰
۴۱۶	۰/۳۵۸۸۲۴	۰/۸۱۰۳۴۴۸۲۸	۰/۸۱۶۷۸۵۷۲۶	۰
۴۱۷	۰/۳۶۴۷۰۶	۰/۲۴۱۳۷۹۳۱	۰/۹۱۹۶۲۴۹۹۷	۱
۴۱۸	۰/۴۲۳۵۲۹	۰	۰/۹۸۰۸۳۹۷۲۹	۰
۴۱۹	۰/۴۹۴۱۱۸	۰/۳۷۹۳۱۰۳۴۵	۰/۱۱۱۲۶۲۰۵	۰
۴۲۰	۰/۴۴۷۰۵۹	۰/۴۳۱۰۳۴۴۸۳	۰/۶۱۱۸۹۳۰۶۵	۰

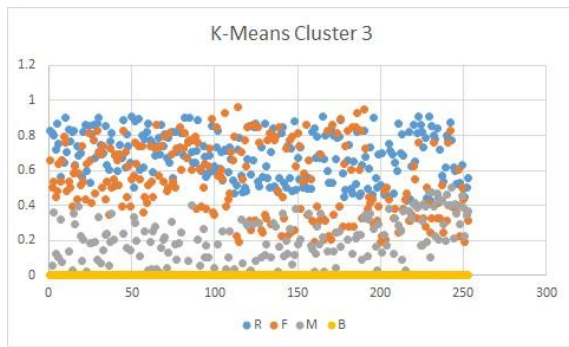
کلیدی بودن یا نبودن مشتری بانک است و در بانک ملت مشتریان کلیدی بانکی در هر استان بنا بر نظر خیرگان بانکی با ویژگی‌های خاصی تعریف می‌گردند، انتخاب شده است.

در جدول (۱) نمونه‌های ده‌تایی از این داده‌های خام که نشان‌دهنده اطلاعات مربوط به مشتریان بانک ملت می‌باشد، نمایش داده شده است.

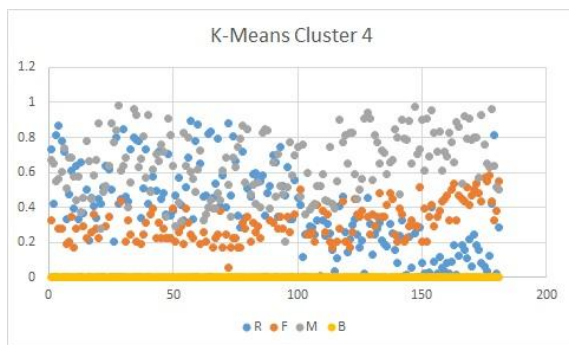
جدول (۱): نمونه ده‌تایی از داده‌های مربوط به مشتریان بانک ملت

ID	Sex	Age	Education	Marital Status	Job	Location	R	F	M	K
۱	مرد	۵۴	کارشناسی	متاهل	فرهنگی	بابل	۱۳۵	۲۶	۳۴۰۰۰۰۰	بله
۲	مرد	۶۹	سیکل	متاهل	آزاد	تهران	۱۵	۲	۶۹۹۵۲۱	بله
۳	مرد	۴۹	دپلم	متاهل	کشاورز	قم	۳۱۱	۷	۲۹۹۳۱۹۹۲۵	بله
۴	مرد	۳۳	کارشناسی ارشد	مجرد	آزاد	تهران	۴۸	۲۹	۸۲۸۰۱۰۰۰۰	بله
۵	مرد	۴۱	دپلم	متاهل	آزاد	تهران	۲۰	۳۹	۳۳۵۳۳۱۱۵	بله
۶	زن	۳۴	کارشناسی	متاهل	کارمند	کرمان	۶۵	۱۱	۶۶۳۰۹۷۲۸	بله
۷	مرد	۴۱	دپلم	متاهل	آزاد	تبریز	۱۸	۱۵۰	۲۲۴۶۸۰۳۱۷۶	بله
۸	مرد	۴۴	دپلم	متاهل	آزاد	تهران	۷	۲۱	۵۲۰۵۲۱۳۹	بله
۹	مرد	۴۳	کارشناسی	متاهل	اتوبوس، نمایشگاه	اصفهان	۲	۸۰	۵۱۲۹۶۵۰۱۵	بله
۱۰	زن	۶۷	بی‌سواد	متاهل	خانه‌دار	اصفهان	۳۴	۶	۸۵۶۶۲۵۴۳	بله

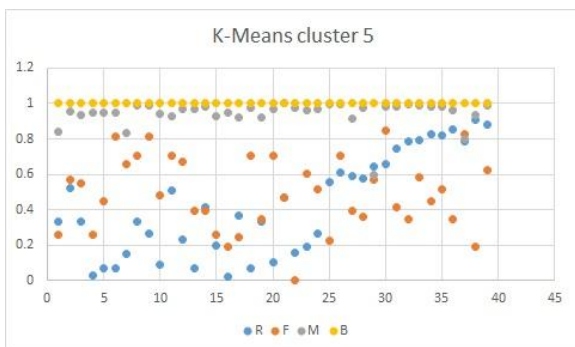
در این تحقیق به جهت حفظ اطلاعات مشتریان بانکی شماره شناسایی مشتریان با شماره‌های قراردادی شماره‌گذاری شده



شکل (۴): خوشه سوم، الگوریتم K-Means



شکل (۵): خوشه چهارم، الگوریتم K-Means



شکل (۶): خوشه پنجم، الگوریتم K-Means

۲-۳-۵- ارزیابی خوشه‌ها به روش مجموع مربع خطاها و انتخاب بهترین روش

در این گام لازم است مجموع مربع خطاهای خوشه‌های حاصل از اعمال الگوریتم A-H-Means با $Q=5$ تا $Q=-5$ محاسبه شود تا مقدار Q برای حالتی که کمترین مجموع مربع خطاها را نتیجه می‌دهد به‌عنوان روش منتخب گزیده شود و خوشه‌های حاصل از روش مذکور جهت استخراج دانش در گام بعد مورد مطالعه قرار گیرند.

$$SSE = \sum_{i=1}^K \sum_{x \in C_i} \sum_{j=1}^n (m_j^i - x_j)^2 \quad (۶)$$

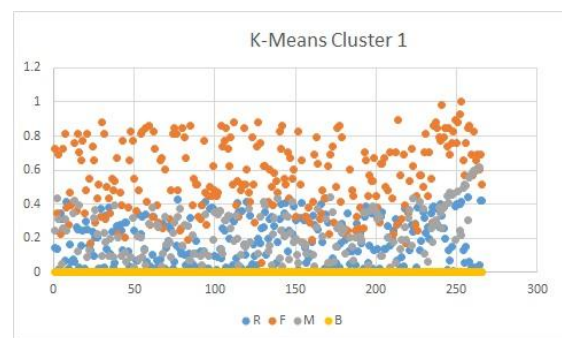
۲-۳-۳- گام تعیین تعداد بهینه خوشه‌ها

تعداد بهینه خوشه‌ها می‌تواند با توجه به اهداف تحقیق و دیدگاه‌های مختلف خبرگان متفاوت باشد. در این تحقیق با استفاده از مجموعه داده‌های مورد استفاده و بررسی نظرات خبرگان بانک ملت بهترین گزینه برای تعیین تعداد مناسب خوشه‌ها، جهت تفکیک رفتاری مشتریان این بانک عدد ۵ به دست آمد. این عدد به‌عنوان یک ورودی در الگوریتم‌های خوشه‌بندی مورد نظر به کار گرفته خواهد شد.

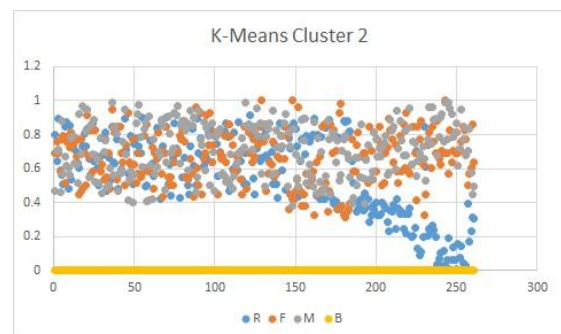
۲-۳-۴- گام خوشه‌بندی به روش A-H-Means

الگوریتم پیشنهادی جهت خوشه‌بندی داده‌های این تحقیق است که جزئیات این روش پیش از این به تفصیل مطرح شد. در این کاربرد $Q \in [-5,5]$ انتخاب شده است که در این صورت $Q = 0$ میانگین حسابی و اگر $Q = -1$ میانگین هارمونیک را نتیجه خواهد داد.

شکل‌های ۲ تا ۶ نمودارهای مربوط به خوشه‌های حاصل از الگوریتم A-H-Means با مقدار $Q=0$ را نشان می‌دهد.



شکل (۲): خوشه اول، الگوریتم K-Means



شکل (۳): خوشه دوم، الگوریتم K-Means

جدول (۴): اطلاعات مربوط به خوشه‌بندی مشتریان بانک ملت به

روش A-H-Means

K	میانگین متغیرها در خوشه			تعداد عناصر خوشه	خوشه
	M	F	R		
۰	۰/۱۷۱۹۷۳	۰/۵۲۸۳۷۲	۰/۱۸۶۲	۲۶۶	خوشه ۱
۰	۰/۷۱۱۵۲۹	۰/۶۶۹۶۳۹	۰/۵۵۹۲۷	۲۶۱	خوشه ۲
۰	۰/۱۴۹۷۷۶	۰/۵۲۲۳۵۲	۰/۶۸۰۲۴	۲۵۳	خوشه ۳
۰	۰/۶۴۹۲۷۲	۰/۲۵۹۷۶۴	۰/۳۸۷۰۳	۱۸۱	خوشه ۴
۱	۰/۹۴۴۴۷۵	۰/۴۹۱۱۵۸	۰/۴۲۶۰۹	۳۹	خوشه ۵

۳- نتایج و بحث

با بررسی جدول (۴) و بهره‌گیری از نظرات خبرگان بانکی می‌توان دریافت که خوشه اول که مقدار متغیر R آن پایین (یعنی اخیراً به بانک مراجعه داشته‌اند)، متغیر F یعنی تعداد تراکنش اخیر آن‌ها طی یک سال گذشته آن نسبتاً بالا و متغیر M یعنی میانگین مبلغ تراکنش‌های آن‌ها در طی یک سال گذشته نسبتاً پایین است، مشتریانی هستند وفادار با مانده حساب پایین (مشتریان قدیمی‌تر) و بانک در مواجهه با این گروه باید به دنبال جواب این سؤال باشد که چطور می‌توان این گروه از مشتریان را که وفاداری نسبتاً بالایی به بانک دارند به سرمایه‌گذاری بیشتر و افزایش موجودی و فعالیت حساب از طرق مختلف در بانک ترغیب نمود. ارائه تسهیلات، می‌تواند پیشنهادی مناسب جهت نگهداشت مشتریان این گروه باشد.

خوشه دوم که مقدار R، F و M نسبتاً بالایی دارند، از آنجایی که میانگین مبلغ تراکنش‌ها و تعداد تراکنش بالایی دارند، مشتریانی ارزشمند برای بانک محسوب می‌شوند. اما از آنجاکه فاصله زمانی آخرین تراکنش آن‌ها با بانک عدد به نسبت بزرگی است، احتمال فرار این مشتریان بالاست. بانک باید با ارائه خدمت‌رسانی سریع‌تر و بهتر به این دسته از مشتریان، این احتمال را کاهش داده و جهت جلب اعتماد مجدد آن‌ها تلاش نماید. در مورد مشتریان خوشه سوم می‌توان گفت این مشتریان با وجود تعداد گردش بالا در یک سال گذشته، اخیراً گردش حسابی نداشته‌اند. در مورد مشتریان این خوشه و خوشه دوم می‌توان گفت باید به روش‌هایی مانند معرفی طرح‌های جدید به این مشتریان سعی در جذب مجدد آن‌ها نمود. البته مشتریان خوشه دوم به دلیل داشتن

مادهای به کار رفته در رابطه (۶) به شرح زیر می‌باشد:

X: یک نقطه در خوشه C_i

J: آمین ویژگی X

J: آمین ویژگی نقطه نماینده خوشه C_i

می‌توان نشان داد که m_{ij} با مرکز خوشه متناسب است.

با توجه مقادیر SSE حاصل در جدول (۳) در این مرحله A-H-Means با مقدار $Q=0$ که معادل روش K-Means می‌باشد بهترین تناسب را با داده‌های مورد نظر داشته‌اند.

جدول (۳): مقادیر مجموع مربع خطاها در الگوریتم‌های مختلف

خوشه‌بندی

SSE	الگوریتم	A-H-Means
۵۷۸/۸۴۶۳۳۹۶۵۸۴۷۸	Q= -۵	
۴۵۱/۹۹۸۶۵۶۶۲۳۸۵۲	Q= -۴	
۴۱۳/۶۳۳۹۵۱۸۷۲۵۳۹	Q= -۳	
۴۵۲/۳۰۱۲۶۹۲۱۰۵۸۲	Q= -۲	
۴۲۹/۵۳۴۷۹۴۰۲۰۳۱۲	Q= -۱	
۹۹/۹۸۴۳۶۲۴۱۷۱۵۲۴	Q= ۰	
۲۶۳/۰۳۲۹۵۴۲۹۷۷۴۹	Q= ۱	
۳۸۷/۲۹۶۸۱۵۵۰۰۴۳۴	Q= ۲	
۴۶۱/۳۷۱۴۰۲۵۱۶۸۴	Q= ۳	
۵۲۰/۴۳۱۷۶۶۰۳۰۳۳۶	Q= ۴	
۵۶۳/۳۴۳۱۹۸۱۹۹۵۶۸	Q= ۵	

۲-۳-۶- گام به‌کارگیری دانش حاصل از خوشه‌بندی

پس از این که مناسب‌ترین الگوریتم برای خوشه‌بندی داده‌های مربوط به مشتریان بانک ملت شناسایی شده و مشتریان در پنج خوشه بخش‌بندی شدند نوبت به بهره‌گیری از دانش حاصل از این خوشه‌ها می‌رسد. در این مرحله با حضور خبرگان بانکی و با مراجعه به پروفایل مشتریان بانک می‌توان به بررسی و تحلیل خوشه‌ها پرداخت و از آنجا که به‌طور معمول مشتریانی که در یک خوشه قرار می‌گیرند دارای ویژگی‌های رفتاری مشابهی خواهند بود، دانش استخراج شده می‌تواند راهنمای بسیار خوبی جهت اتخاذ راهبردهای بهینه متناسب با هر خوشه برای مدیریت هر چه بهتر ارتباط با مشتریان باشد. جدول (۴) بخشی از اطلاعات حاصل از خوشه‌بندی نمونه هزارتایی از داده‌های مربوط به مشتریان بانک ملت را نشان می‌دهد.

ویژگی‌های اکثریت مشتریان هر گروه، سرویس‌های متناسب با این ویژگی‌ها را به روش‌های متفاوت به مشتریان پیشنهاد نمود.

می‌توان با بررسی خوشه‌ها، خوشه‌هایی که بیش‌ترین تعداد مشتریان کلیدی در آن‌ها قرار گرفته‌اند شناسایی نمود. با این امکان مشتریان مستعد کلیدی شدن شناخته می‌شوند و با خدمت‌رسانی هر چه بیشتر به این گروه‌ها منجر به حفظ و ادامه ارتباط مؤثر آن‌ها با بانک گردید.

با بررسی خوشه‌هایی که بیش‌ترین مشتریان کلیدی را در خود جای داده‌اند و شناسایی اکثریت مشتریان در این خوشه‌ها از منظر نوع شغل، محل منطقه بانکی و جنسیت و ... و با نظر کارشناسان و خبرگان بانکی می‌توان خدمات بانکی ویژه و متناسب با این گروه‌ها را طراحی نموده و از طریق پیامک و یا به‌طور رودررو به این افراد معرفی نمود.

۵- مراجع

- [1] H. Hassani, X. Huang, and E. Silva, "Digitalisation and Big Data Mining in Banking," *Big Data Cogn. Comput.*, pp. 2-18 2018.
- [2] H. Mortazi and S. Parsa, "Detecting Baht Network with Behavioral Analysis of Network Flow and Using Data Mining Algorithms," *Journal of Electronical & Cyber Defence*, vol. 5, no. 3, pp. 1-15, 2018. (In Persian)
- [3] M. Garcia-Murillo and H. Annabi, "Customer knowledge management," *Journal of the Operational Research Society*, pp. 875-884, 2002.
- [4] E. W. T. Ngai, Li Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Systems with Applications*, vol. 36.2, pp. 2592-2602, 2009.
- [5] J. Ranjan, "A Review of Data Mining Tools In Customer Relationship Management," *Journal of Knowledge Management Practice*, vol. 9, no. 1, March 2008.
- [6] N.-C. Hsieh, "An integrated data mining and behavioral scoring model for analyzing bank customers," *Expert Systems with Applications*, vol. 27.4, pp. 623-63, 2004.
- [7] F.-L. Chen, et al., "Combination of feature selection approaches with SVM in credit scoring," *Expert Systems with Applications*, vol. 37, pp. 4902-4909, 2010.
- [8] S. M. Hosseini, A. Maleki, and M. R. Gholamian, "Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty," *Expert Systems with Applications*, vol. 37.7, pp. 5259-5264, 2010.
- [9] D. Dong, J. Zhang, and J. Ye, "Research on Customer Segmentation Method of Commercial Bank Based on Data Mining," *3rd International Conference on Innovation Development of E-commerce and Logistics*, pp. 62-65, 2017.
- [10] Y.-P. XI and C. H. E. N. Min, "Application of data mining technology in CRM system of commercial banks," *DEStech*

میانگین مبلغ بالاتر نسبت به مشتریان این خوشه از ارزش بالاتری برخوردارند.

مشتریان خوشه سوم مشتریانی نسبتاً ارزشمند ولی با فعالیت حساب پایین و به‌طور معمول مشتریان جدید هستند که میزان نسبتاً بالای تبادل نقدینگی با بانک را دارا می‌باشند. بانک می‌تواند با معرفی کامل کلیه خدمات بانکی ارائه شده به روش‌های مختلف مانند ارائه کاتالوگ، ارسال پیامک، ایمیل و ... مشتری جدید را در انتخاب خدمات متناسب و در نتیجه بقای ارتباط با این بانک یاری نماید.

مقدار میانگین ۱ مربوط به متغیر K در خوشه پنجم نشان می‌دهد کلیه مشتریان این خوشه جزء مشتریان کلیدی بانک می‌باشد. با گسترش دامنه مطالعه به سایر متغیرهای مربوط به این مشتریان می‌توان به گزاره‌های دانشی ارزشمندی دست یافت. مثلاً اکثریت افراد این خوشه مرد، متأهل و دارای شغل آزاد می‌باشند. این دانش در مورد ارزشمندترین مشتریان بانک می‌تواند در زمینه اخذ تصمیماتی جهت افزایش روزافزون رضایت این دسته از مشتریان و ممانعت از رویگردانی آن‌ها از این بانک بسیار کمک کننده باشد. یکی از این رویکردها که می‌تواند در ارتقاء رضایت این مشتریان تأثیرگذار باشد، ارسال پیامک‌هایی خاص این گروه‌ها و معرفی خدمات جدید به قصد افزایش بازدهی حساب‌های آنان در نزد این بانک می‌باشد. توسعه خدمات جدید متناسب با مطالبات این گروه می‌تواند موجب بقای ارتباط آن‌ها با بانک بوده و قطعاً رضایت مشتری در درآمد سودآوری زیادی را برای بانک به همراه خواهد داشت.

بررسی نتایج حاصل از بکارگیری این راهبردها در بانک ملت نشان می‌دهد که قابلیت سازمان برای جلب وفاداری مشتریان خود با راهبردهایی که با توجه به این دسته‌بندی به کار برده است در مقایسه با روش معمول انتخاب تصادفی که در اکثر سازمان‌ها در ایران به کار می‌رود، افزایش چشمگیری داشته است.

۴- نتیجه‌گیری

با به‌کارگیری روش‌های هوشمند داده‌کاوی مشتریان بانک ملت در پنج دسته و با توجه به بیشترین شباهت در الگوی رفتاری گروه‌بندی شدند.

خوشه‌بندی مشتریان بر مبنای چارچوب ارائه شده در این تحقیق به ما کمک می‌کند تا مشتریانی که دارای الگوهای رفتاری مشابه برحسب متغیرهای تراکنشی مشتریان و متغیرهای جمعیت آماری هستند را در گروه‌های واحدی در نظر گرفته و بر مبنای

- [13] Z. Rezvani, "New Product Development Based on Customer Knowledge Management," Master Thesis, Luleå University of Technology, pp. 25-35, 2009.
- [14] S. M. Hosseini Bamkan, "Application of Data Mining Techniques for Improving Customer Relationship Management in the Banking Industry (Case Study: Parsian Bank)," Master's Thesis, Supervisor: Dr. Mohammad Reza Taghva, Faculty of Accounting and Management, University Allameh Tabatabai, 2009. (In Persian)
- Transactions on Engineering and Technology Research eeta, 2017.
- [11] H. R. Momeni, et al., "Using data mining technology in customer relationship management in the banking system," National Conference on Information Technology and Jihad-e-Economics, Salman Farsi University, Kazeroun, March 2011. (In Persian)
- [12] J. Han, J. Pei, and M. Kamber, "Data mining: concepts and techniques," Elsevier, 2011.