

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

أَمَّنْ هُوَ قَانِتٌ آنَاءَ اللَّيْلِ سَاجِداً وَ قَائِماً يَحْذَرُ الْآخِرَةَ وَ يَرْجُوا رَحْمَةَ رَبِّهِ **قُلْ**
هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَ الَّذِينَ لَا يَعْلَمُونَ إِنَّمَا يَتَذَكَّرُ أُولُوا الْأَلْبَابِ

[آیا چنین کسی با ارزش است] یا کسی که در ساعات شب به عبادت مشغول است و در حال سجده و قیام، از عذاب آخرت می ترسد و به رحمت پروردگارش امیدوار است؟! **بگو: (آیا کسانی که می دانند با کسانی که نمی دانند یکسانند؟! تنها خردمندان متذکر می شوند!)**

(سوره ی مبارکه الزمر/ ۹)

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



وزارت علوم، تحقیقات و فناوری
دانشگاه علوم و فنون مازندران
پایان نامه مقطع کارشناسی ارشد
رشته مهندسی فناوری اطلاعات

مدیریت دانش مشتریان بانک مهر اقتصاد با استفاده از تکنیک‌های داده‌کاوی



اساتید راهنما: دکتر جواد وحیدی، دکتر بابک شیرازی

استاد مشاور: دکتر حسین علیزاده

استاد داور: پروفسور ایرج مهدوی

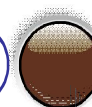
دانشجو: سیده عطیه علی‌کیا امیری

(89191119, a.alikia@ustmb.ac.ir)

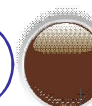
فهرست مطالب



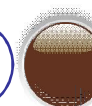
مقدمه و کلیات تحقیق



مروری بر پیشینه



روش تحقیق

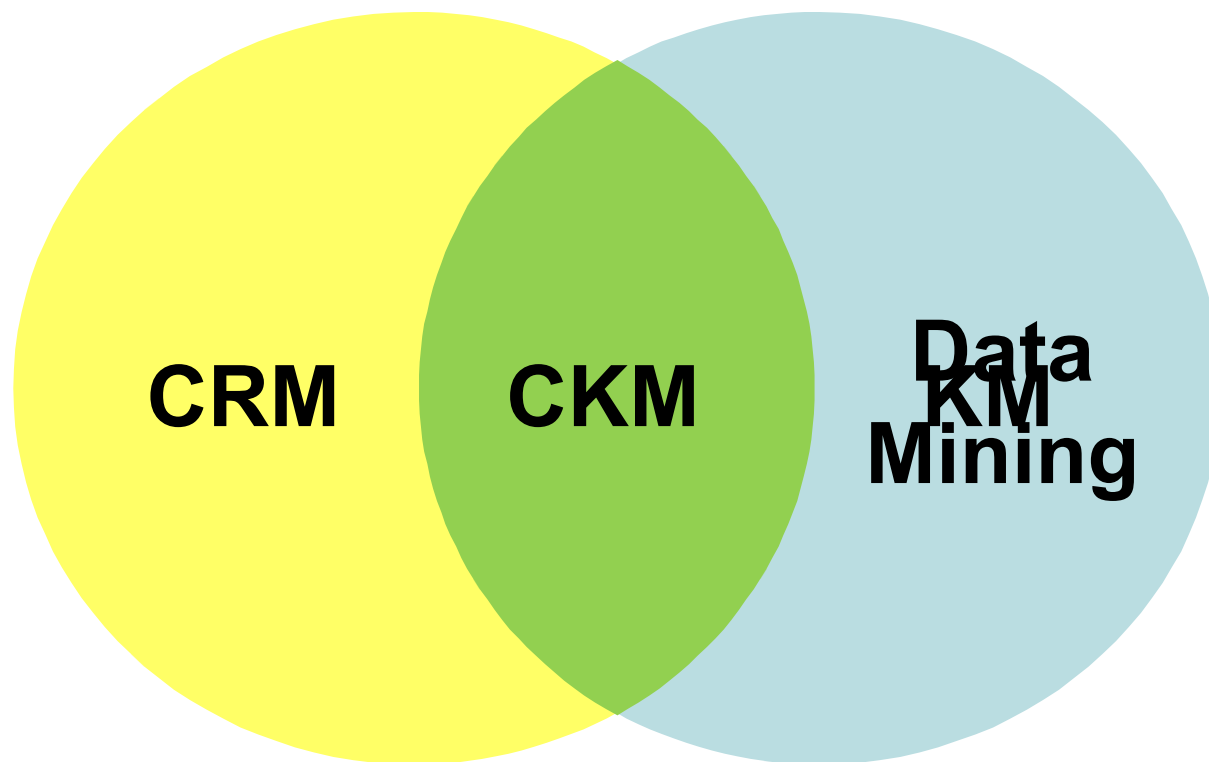


نتیجه گیری



پیشنهادات





افزایش روزافزون سطح رقابت در بازار

اهمیت یافتن مشتری در عرصه پرتلاطم رقابت بانک‌ها

ورود فناوری‌های جدید به سازمان‌ها و افزایش سرعت تولید اطلاعات

حجم بالای پایگاه‌های داده و پراکندگی و عدم به‌کارگیری راهکارهای مناسب جهت تحلیل این داده‌ها

اهمیت و
ضرورت
انجام
تحقیق

اهداف تحقیق

هدف اصلی

- تحلیل داده‌های حاصل از تعامل بانک با مشتریان، با هدف شناسایی مشخصه‌های هر گروه با به‌کارگیری نرم‌افزارهای موجود در زمینه داده‌کاوی و ابزارهایی از آن نظیر خوشه‌بندی.

اهداف فرعی

- عرضه‌ی محصول طراحی شده متناسب با نیازهای هر گروه
- عرضه محرک‌های مناسب برای توسعه ارتباط مشتری با بانک
- تولید محصول جدید متناسب با ویژگی هر گروه
- حصول هر چه بیشتر رضایت مشتری

تحقیقات مرتبط

تحقیقات مروری و دسته‌بندی (Ngai, 2009)

کاربرد داده‌کاوی در ارزیابی اعتبار مشتریان (Lee, Byungtae and et al, 2010)

کاربرد داده‌کاوی در زمینه کشف تقلب (Chen, Fei-Long and et al, 2010)

کاربرد داده‌کاوی در تحلیل روی‌گردانی مشتری (Nie, Guangli, et al, 2011)

کاربرد داده‌کاوی در بخش‌بندی و مدل‌سازی رفتاری مشتریان در صنعت بانکداری

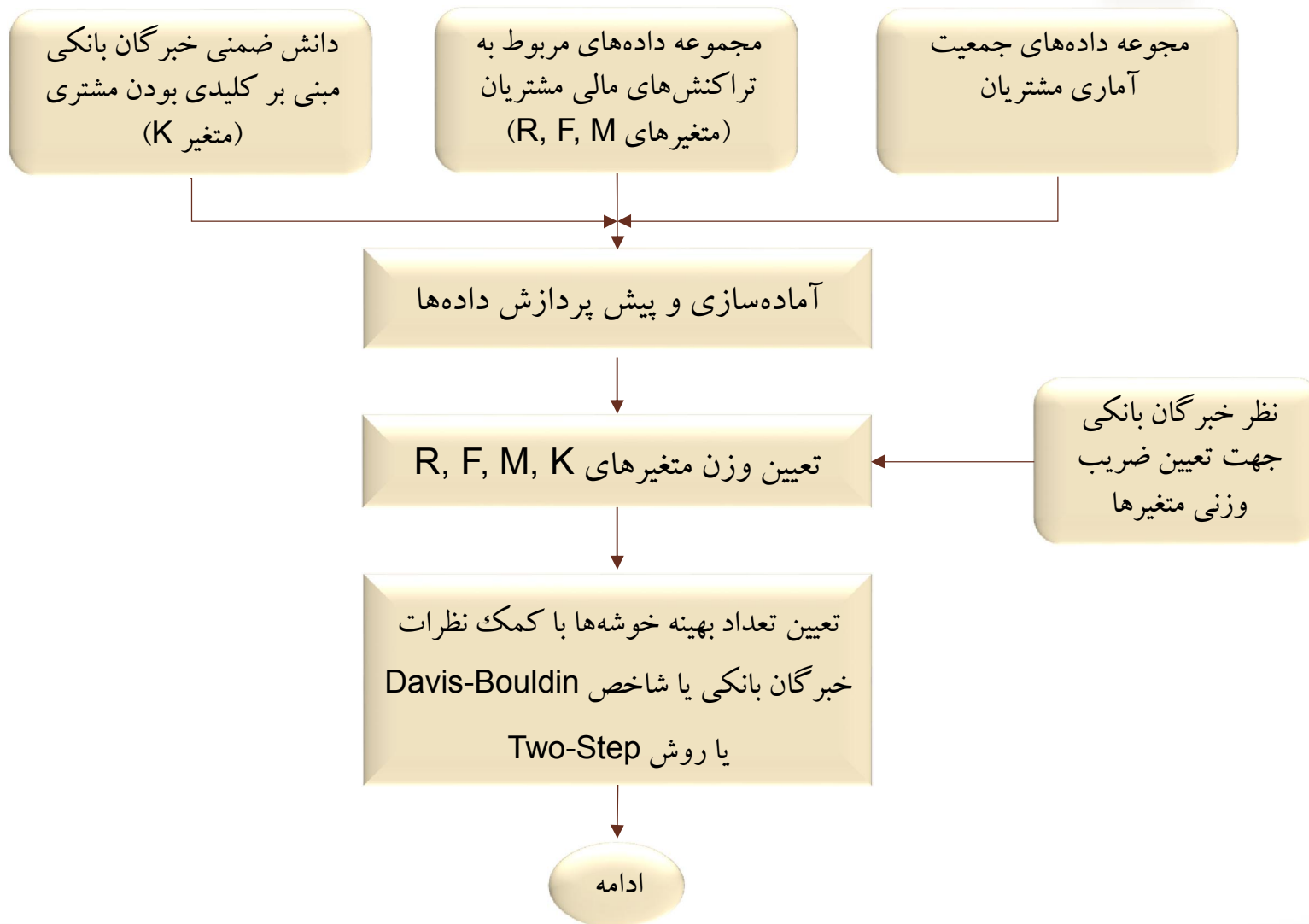
کاربرد
داده‌کاوی
در
صنعت
بانکداری

تحقیقات مرتبط

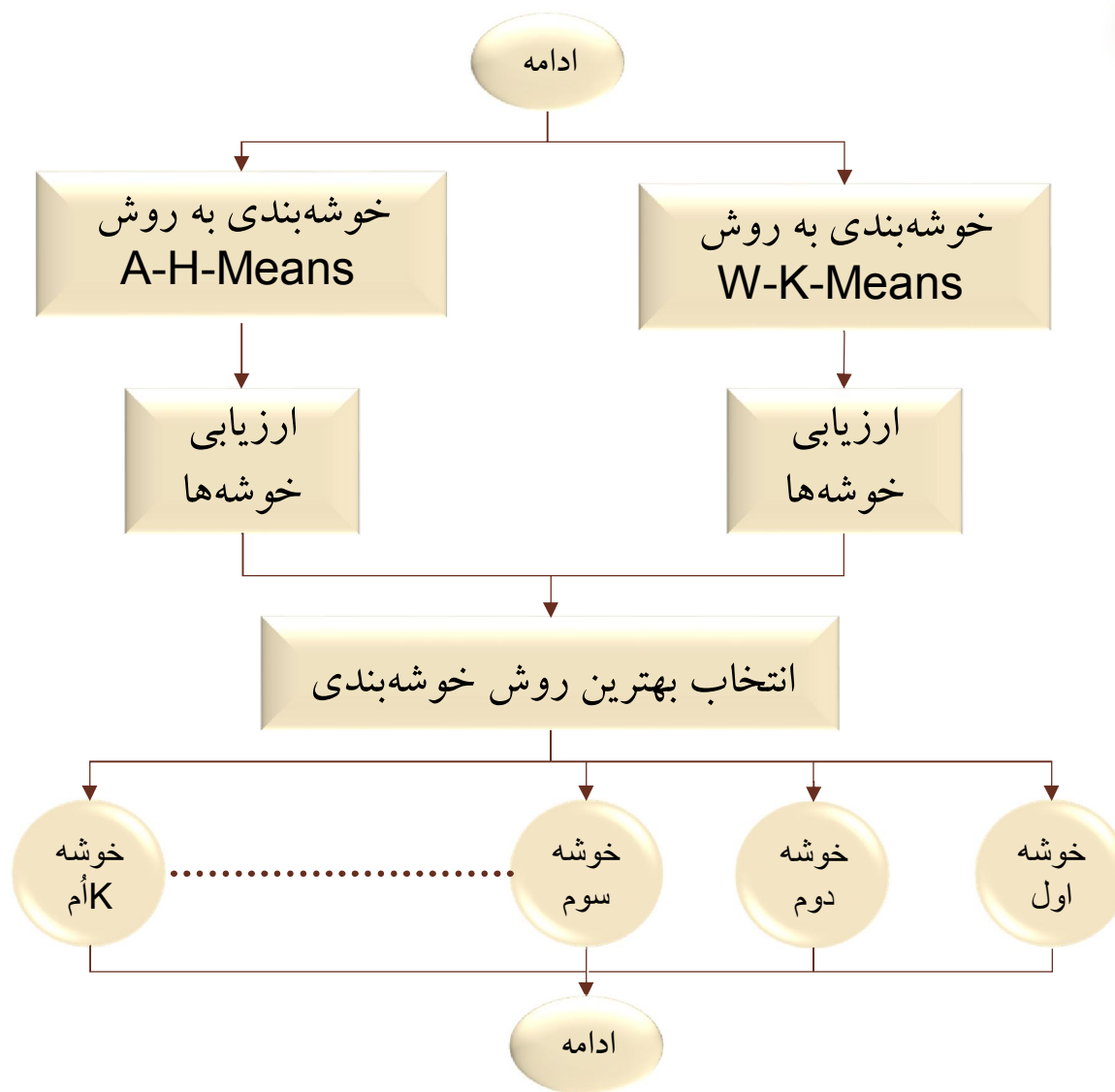
کاربرد داده کاوی در بخش بندی و مدل سازی رفتاری مشتریان در صنعت بانکداری

مؤلف	هدف	روش های مورد استفاده	ورودی ها
۲۰۱۰ SMS Hosseini, et al	بخش بندی مشتریان جهت تخمین وفاداری مشتری	داده های جمعیت شناختی، داده های وزن دار مربوط به تراکنش های مالی (WRFM)	خوشه بندی با استفاده از روش K-means و SOM
۱۳۹۰ مؤمنی و همکاران	دسته بندی مشتریان جهت شناسایی مشتریان طلایی طبق اصل ۲۰/۸۰ پارتو	اطلاعات مربوط به مشتریان وام گیرنده از بانک	شبکه های عصبی
۱۳۸۸ حسینی، سید مجتبی	بخش بندی مشتریان جهت مدیریت ارتباط با آنان و کشف سرویس های مرتبط با هر بخش	متغیرهای جمعیت شناختی، متغیرهای RFM مشتریان بانک و لیستی از خدمات ارائه شده توسط بانک. مورد مطالعه: بانک پارسیان	شبکه های عصبی، خوشه بندی توسط تکنیک های خود سازمان دهنده (SOM) و الگوریتم Apriori برای کشف وابستگی بین خدمات ارائه شده توسط بانک با مشتریان.

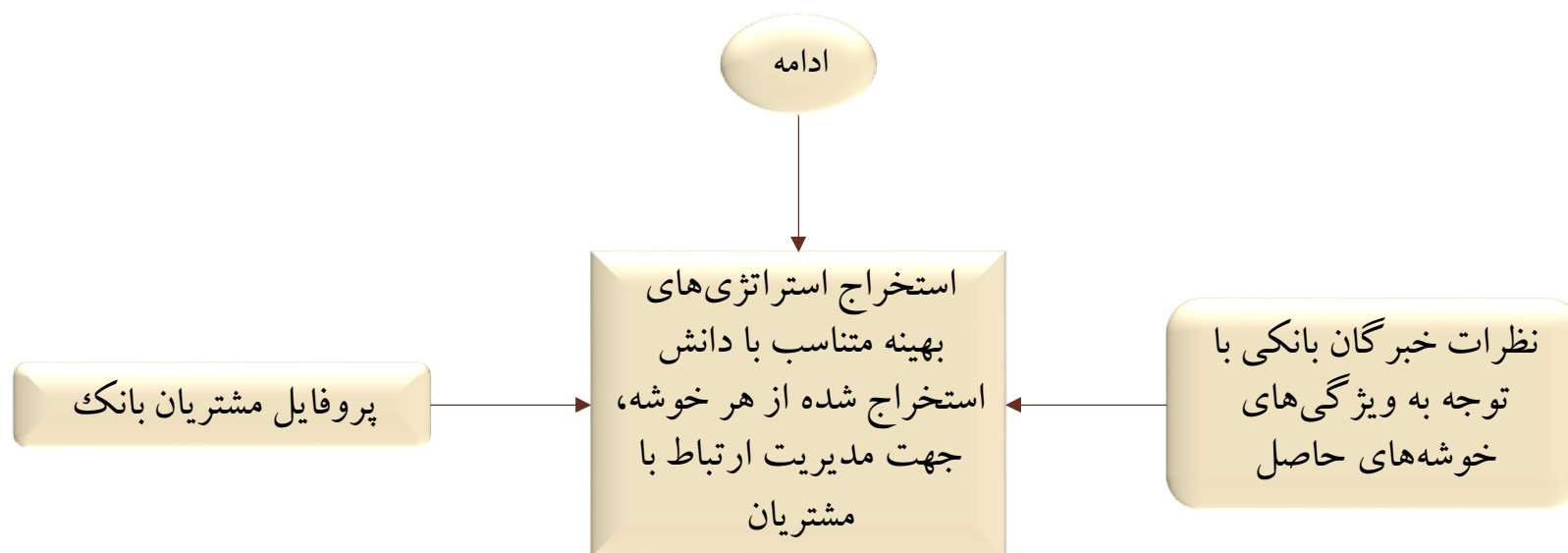
روش تحقیق: چارچوب پیشنهادی



چارچوب تحقیق (ادامه)



چارچوب تحقیق (ادامه)



• بانک مهر اقتصاد

- مؤسسه قرض الحسنه بسیجیان سابق
- هفتصد شعبه تحت پوشش
- بزرگترین شعبه اتوماسیون در بین بانک‌های خصوصی

• موضوع فعالیت:

- حضور موثر در عرصه‌های اقتصادی (پولی و بانکی)
- ارائه خدمات گوناگونی در حمایت از پروژه‌های مولد و اشتغال‌زا، ازدواج، درمان، تحصیل و ...
- ارائه خدمات الکترونیک
- ارائه خدماتی چون افتتاح انواع سپرده‌های سرمایه‌گذاری و سود مناسب، اعطای تسهیلات در بخش‌های صنعتی و تولیدی، مسکن، خرید خودرو و لوازم خانگی و خدمات اعتباری.
- رویکرد بانکداری اسلامی

گام اول: انتخاب متغیرها



- داده‌های جمعیت آماری مشتریان
 - شامل جنسیت، سن، تحصیلات، وضعیت تأهل، شغل، محل منطقه بانکی.
- تراکنش‌های مالی مشتریان
 - تراکنش اخیر مشتری (Recency)
 - متغیر پولی (Monetary)
 - فراوانی (Frequency)
- دانش ضمنی خبرگان بانکی در مورد کلیدی بودن مشتری
 - متغیر K

نمونه ده‌تایی از داده‌های مربوط به مشتریان بانک مهر اقتصاد

ID	Sex	Age	Education	Marital Status	Job	Location	R	F	M	K
1	مرد	54	کارشناسی	متاهل	فرهنگی	بابل	135	26	3400000	خیر
2	مرد	69	سیکل	متاهل	آزاد	بابل	15	2	699521	خیر
3	مرد	29	دیپلم	متاهل	کشاورز	بابل	311	7	299319925	خیر
4	مرد	23	کارشناسی ارشد	مجرد	آزاد	بابل	48	29	828010000	خیر
5	مرد	41	دیپلم	متاهل	آزاد	بابل	20	39	234543115	خیر
6	زن	34	کارشناسی	متاهل	کارمند	بابل	65	11	66309728	خیر
7	مرد	41	دیپلم	متاهل	آزاد	بابل	18	150	22468073176	خیر
8	مرد	44	دیپلم	متاهل	آزاد	بابل	7	21	52052139	خیر
9	مرد	43	کارشناسی	متاهل	نمایشگاه اتومبیل	بابل	2	80	5129665015	بله
10	زن	67	بی‌سواد	متاهل	خانه‌دار	بابل	34	6	85662543	خیر

گام دوم: آماده‌سازی و پیش‌پردازش داده‌ها



1. استخراج نمونه هزار تایی از داده‌های موجود در پایگاه اطلاعاتی بانک مهر اقتصاد
2. انتخاب متغیرهای R ، F ، M و K به عنوان ورودی الگوریتم‌های خوشه‌بندی.
3. نرمال سازی داده‌ها به روش Min-Max

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

در فرمول بالا X_i نشان‌دهنده i امین مقدار متغیر X و X_i^* مقدار نرمال شده این متغیر است.

نمونه ده‌تایی از متغیرهای نرمال شده

K	M	F	R	ID
0	0.394722852	0	0.035294	411
0	0.46601071	0.689655172	0.8	412
0	0.770616295	0	0.205882	413
0	0.90193016	0.620689655	0.1	414
0	0.213265911	0	0.070588	415
0	0.816785726	0.810344828	0.358824	416
1	0.919624997	0.24137931	0.364706	417
0	0.980839729	0	0.423529	418
0	0.11126205	0.379310345	0.494118	419
0	0.611893065	0.431034483	0.447059	420

گام سوم: تعیین تعداد بهینه خوشه‌ها



با استفاده از مجموعه داده‌های مورد استفاده و بررسی نظرات خبرگان بانک مهر اقتصاد بهترین گزینه برای تعیین تعداد مناسب خوشه‌ها، جهت تفکیک رفتاری مشتریان این بانک عدد ۵ به دست آمد.

گام چهارم: خوشه‌بندی W-K-Means خوشه‌بندی به روش



۱- الگوریتم با انتخاب ۵ مرکز آغاز می‌شود. این مراکز می‌تواند به طور حذسی یا تصادفی انتخاب گردد.

۲- محاسبه مقدار تابع هدف به صورت زیر است:

$$KM(X, C) = \sum_{i=1}^n \min_{j \in \{1 \dots k\}} \|x_i - c_j\|^2$$

۳- برای هر داده x_i ، عضویت $m(c_j|x_i)$ به ازای هر مرکز c_j محاسبه می‌شود.

تابع عضویت به صورت زیر محاسبه می‌شود:

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

گام چهارم: خوشه‌بندی خوشه‌بندی به روش W-K-Means (ادامه)



۴- برای هر دسته، مجدداً مرکز c_j جدید به صورت زیر محاسبه می‌شود:

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

- $X = \{x_1, x_2, \dots, x_n\}$: داده‌ای که باید خوشه‌بندی شود.

- $C = \{c_1, c_2, \dots, c_K\}$: مجموعه مراکز خوشه‌ها.

- $w(x_i)$ وزن داده x_i .

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

گام چهارم: خوشه‌بندی خوشه‌بندی به روش W-K-Means (ادامه)



- در اکثر روش‌های داده‌کاوی ارزش همه داده‌ها یکسان در نظر گرفته می‌شود که این مسئله یکی از نقاط ضعف این الگوریتم‌ها شناخته می‌شود. در روش WK-Means سعی شده با وزن دهی متغیرها بر اساس اهمیت نسبی آن‌ها این ضعف را تعدیل نماید.

متغیر	R	F	M	K
ضریب وزنی	0.12	0.22	0.32	0.34

گام چهارم: خوشه‌بندی A-H-Means خوشه‌بندی به روش

در این روش میانگین جدید هر خوشه توسط فرمول زیر محاسبه می‌گردد:

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

- $X = \{x_1, x_2, \dots, x_n\}$: داده‌ای که باید خوشه‌بندی شود.
- $C = \{c_1, c_2, \dots, c_K\}$: مجموعه مراکز خوشه‌ها.
- Q : متغیر توان که می‌تواند از بازه R انتخاب می‌شود (در این کاربرد $[-5, 5]$).

گام چهارم: خوشه‌بندی
خوشه‌بندی به روش A-H-Means (ادامه)



- در فرمول میانگین آنتی هارمونیک در صورتی که $Q=0$ فرمول

$$\text{میانگین حسابی} \quad \frac{\sum_{i=1}^n \frac{\|x_i - c_j\|}{j \in \{1 \dots k\}}}{n} \quad \text{و در نتیجه روش K-}$$

Means حاصل خواهد شد و اگر $Q=-1$ میانگین **هارمونیک**

و در نتیجه روش KH-Means را نتیجه خواهد داد.

گام پنجم: ارزیابی خوشه‌ها



ارزیابی خوشه‌ها با معیار مجموع مربع خطاها:

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

x : یک نقطه در خوشه C_i

x_j : ژامین ویژگی x

m_j^i : ژامین ویژگی نقطه نماینده خوشه C_i

نتایج ارزیابی خوشه‌ها

SSE	الگوریتم	
99.9843624171528	K-Means	
101.70396570802	W-K-Means	
578.846339658478	Q= -5	A-H-Means
451.998656623852	Q= -4	
413.633951872539	Q= -3	
452.301269210582	Q= -2	
429.534794020312	Q= -1 (H-Means)	
99.9843624171524	Q= 0 (K-Means)	
263.032954297749	Q= 1	
387.296815500434	Q= 2	
461.37140251684	Q= 3	
520.431766030336	Q= 4	
563.343198199568	Q= 5	

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی



میانگین متغیرها در خوشه				تعداد عناصر خوشه	
K	M	F	R		
0	0.171973	0.538372	0.1862	266	خوشه اول
0	0.711529	0.669639	0.55927	261	خوشه دوم
0	0.149776	0.522352	0.68024	253	خوشه سوم
0	0.649272	0.259764	0.38703	181	خوشه چهارم
1	0.944475	0.491158	0.42609	39	خوشه پنجم

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی (ادامه)

خوشه اول

- متغیرها: مقدار متغیر R پایین، F نسبتاً بالا و متغیر M نسبتاً پایین
- وضعیت گروه: مشتریانی هستند وفادار با مانده حساب پایین (مشتریان قدیمی‌تر)
- راهکار: ترغیب برای سرمایه‌گذاری بیشتر و افزایش موجودی و فعالیت حساب از طرق مختلف در بانک
- پیشنهاد: ارائه تسهیلات، می‌تواند پیشنهادی مناسب جهت نگهداشت مشتریان این گروه باشد.

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی (ادامه)

خوشه دوم

- متغیرها: مقدار R ، F و M نسبتاً بالا
- وضعیت گروه: مشتریانی ارزشمند با احتمال فرار بالا
- راهکار: سعی در کاهش این احتمال و جلب مجدد اعتماد آنها
- پیشنهاد: خدمت رسانی سریع‌تر و بهتر در باجه‌ها، جذب مجدد آنها با معرفی طرح‌های جدید به آنها

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی (ادامه)

خوشه سوم

- متغیرها: مقدار متغیر R بالا، F نسبتاً بالا و متغیر M پایین
- وضعیت گروه: با وجود تعداد گردش بالا در یک سال گذشته، اخیراً گردش حسابی نداشته‌اند.
- راهکار: ترغیب برای سرمایه‌گذاری بیشتر و افزایش موجودی و جذب مجدد
- پیشنهاد: معرفی طرح‌های جدید

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی (ادامه)

خوشه چهارم

- متغیرها: مقدار متغیر R پایین، F پایین و متغیر M نسبتاً بالا
- وضعیت گروه: نسبتاً ارزشمند ولی با فعالیت حساب پایین و به طور معمول مشتریان جدید با میزان نسبتاً بالای تبادل نقدینگی با بانک
- راهکار: کمک به مشتری جدید جهت انتخاب خدمات متناسب و در نتیجه بقای ارتباط با این بانک
- پیشنهاد: معرفی کامل کلیه خدمات بانکی ارائه شده به روش‌های مختلف مانند ارائه کاتالوگ، ارسال پیامک، ایمیل و ...

گام ششم: به کارگیری دانش حاصل از خوشه‌بندی (ادامه)

خوشه پنجم

- متغیرها: مقدار میانگین ۱ مربوط به متغیر K
- وضعیت گروه: مشتریان کلیدی بانک
- راهکار: اخذ تصمیماتی جهت افزایش روزافزون رضایت این دسته از مشتریان
- پیشنهاد: توسعه خدمات جدید متناسب با مطالبات این گروه

نتایج نهایی

- بکارگیری دانش حاصل جهت بسط و ارتقاء نگاه مدیریتی تحت عنوان سیستم پشتیبان تصمیم مدیران بانک برای ارائه خدمات به مشتریان بر خلاف نظام‌های تصمیم‌گیری سنتی برای اولین بار در بین بانک‌های خصوصی ایران.
- بهره‌گیری از دانش ضمنی خبرگان بانکی جهت دستیابی به نتایج کاربردی‌تر.

پیشنهادهای جهت تحقیقات آتی

به کارگیری ابزار داده کاوی در سایر زمینه ها در بانک مهر اقتصاد مانند کشف تقلب، ارزیابی اعتبار مشتریان و پیش بینی توانمندی مشتری در بازپرداخت اقساط تسهیلات، ارزیابی عملکرد بانک و ...

استفاده از الگوریتم های داده کاوی مانند قواعد همبستگی و استخراج قوانینی جهت شناسایی مشتریان کلیدی بانک

تعیین تعداد بهینه خوشه ها استفاده از روش های **Davis Bouldin, Two-Step** و ... در این کاربرد

ارائه سرویس مناسب به مشتریان بالقوه با استفاده از الگوریتم **Apriori** جهت شناسایی هر چه دقیق تر رابطه میان گروه های مختلف مشتریان و سرویس های مورد علاقه آنها در این بانک.

از مسن توبه شما سیاست‌گذارم

مدیریت دانش مشتریان بانک مهر اقتصاد با استفاده از تکنیک‌های داده‌کاوی