

دسته‌بندی اعتباری مشتریان بانکی با تلفیق الگوهای مفهومی، سلسله مراتبی فازی و خوشه‌بندی

بیژن بیدآباد^۱ روح الله محمدی^۲ طاهره نورسته^۳

چکیده

در این مقاله، به منظور رتبه‌بندی اعتباری مشتریان از تلفیق سه الگوی مفهومی Hsieh و تحلیل سلسله مراتبی فازی FAHP و روش خوشه‌بندی K-means استفاده گردید. عوامل اصلی اثرگذار بر رتبه‌بندی مشتریان شامل داده‌های تراکنشی مشتری، رفتار بازپرداخت و متغیرهای RFM تعریف و برای دسته‌بندی اعتباری مشتریان حقوقی بانک توسعه صادرات ایران، ۵ متغیر مزبور از پایگاه داده‌های بانک استخراج و نرمالایزه گردید. وزن هر یک از متغیرها از طریق مصاحبه با کارشناسان بانکی و با استفاده از روش تحلیل سلسله مراتبی فازی محاسبه شد. با استفاده از مقادیر متغیرهای مزبور با استفاده از الگوریتم K-means، تعداد خوشه‌های بهینه مشتریان محاسبه گردید. نهایتاً مشتریان بانک در ۵ گروه اعتباری رتبه‌بندی و ارزش هر خوشه برآورد گردید.

نتایج بدست آمده نشان می‌دهد که رتبه مشتریان به ترتیب بیشترین تاثیر را از متغیرهای تأخر^۴، رفتار بازپرداخت، تراکنش، تکرر^۵ و پولی می‌پذیرد. براساس نتایج بدست آمده، ۵۴٪ مشتریان در خوشه سوم (با ارزش خوشه ۰/۹۵) و خوشه پنجم (با ارزش خوشه ۰/۷۶) مشتریان خوب و بسیار خوب بانک را تشکیل می‌دهند. ریسک اعتباری این دو گروه (بالاخص خوشه سوم) در حداقل مقدار خود قرار دارد. ۳۲٪ مشتریان در خوشه دوم با ارزش خوشه ۰/۵۹ مشتریان متوسط بانک را به لحاظ اعتباری تشکیل می‌دهند. ۱۴٪ مشتریان دو خوشه چهارم و اول را با ارزش خوشه ۰/۴۲ و ۰/۲۶ تشکیل می‌دهند که پریسک‌ترین مشتریان بانک هستند و از احتمال نکول بالایی برخوردارند.

کلیدواژه‌ها: رتبه‌بندی اعتباری و رفتاری، AHP فازی، RFM، خوشه‌بندی K-means

مقدمه

بانکها و موسسات مالی با استفاده از رتبه‌بندی اعتباری با توجه به سوابق متقاضی تسهیلات، احتمال عدم بازپرداخت تسهیلات توسط وی را ارزیابی کرده و امتیازبندی می‌نمایند. مدل‌های رتبه‌بندی رفتاری و اعتباری یکی از کاربردی‌ترین روش‌های آماری و تحقیق در عملیات در بانکداری است (Thomas, 2000). هدف اصلی از مدل‌های

^۱ - دکتر بیژن بیدآباد، استاد اقتصاد و مشاور ارشد بانکداری اسلامی، بانک ملی ایران.

Web: <http://www.bidabad.com>, Email: bijan@bidabad.com, bidabad@yahoo.com

^۲ - Email: r_mohamadi58@yahoo.com

^۳ - بانک توسعه صادرات ایران

^۴ - Recency

^۵ - Frequency

رتبه‌بندی اعتباری و رفتاری، طبقه‌بندی مشتریان به گروه‌های مختلف است (Lancher et al, 1995) و از طریق آنالیز خوشه‌ای قابل حل می‌باشد (Hand, 1981; Johnson & Wichern, 1998; Morrison, 1990). از این روش با بکارگیری داده‌های بانکی، تحلیل‌های خوشه‌ای جهت رتبه‌بندی اعتباری مشتریان جدید با توجه به ویژگی‌هایی همچون سن، وضعیت تاهل و درآمد نیز استفاده شده است (Chen & Huang, 2003). روش رتبه‌بندی در بانک‌ها جهت پیش‌بینی رفتار خرید مشتریان در آینده یا وضعیت اعتباری موجود آنان با استفاده از تحلیل‌های خوشه‌ای بکار برده شده است (Setiono et al, 1998).

الگوی RFM

یکی از روش‌های مطرح در تحلیل ارزشیابی مشتری، الگوی RFM⁶ است که در سال ۱۹۹۴ ارائه شد (Hughes, 1994). در این الگو، تفاوت مشتریان با استفاده از سه متغیر تاخر (تازگی آخرین خرید)، تکرر خرید و ارزش پولی خرید اندازه‌گیری می‌شود. در این الگو R، فاصله زمانی آخرین خرید مشتری تا زمان حال، F، تعداد خریدها در یک دوره زمانی مشخص و M، مبلغ خرید در دوره مورد نظر تعریف می‌شود (Wang, 2010). بررسی‌ها نشان می‌دهد که هر چه R و F بیشتر باشد، احتمال آنکه تراکنش جدیدی با مشتری صورت بگیرد، بیشتر خواهد بود و هر چه M بیشتر باشد، احتمال بازگشت مشتری برای خرید بیشتر است (Wu and Lin, 2005). مطالعات (Newell, 1997) نشان داده است که متغیرهای مدل RFM برای رتبه‌بندی مشتریان بسیار کارا هستند. به و همکاران برای انتخاب روش بازاریابی مستقیم از متغیرهای RFM استفاده نمود و با اضافه کردن دو متغیر زمان اولین خرید و احتمال ریزش، الگوی RFM را بسط می‌دهد (Yeh et al, 2009). این الگو کاربردهای گوناگون دیگری نیز داشته است. جانکرا از این الگو جهت بخش‌بندی مشتریان به منظور تعیین سیاست‌های بهینه بازاریابی استفاده نمود (Jonkera et al, 2004). در مطالعه دیگری از این مدل جهت رتبه‌بندی مشتریان استفاده شده است (Hsieh, 2004) و (غضنفری و همکاران ۱۳۸۹). برخی نیز از این مدل جهت محاسبه ارزش دوره عمر مشتری⁷ استفاده کرده‌اند (رمزی و قنبری ۱۳۸۸) و (Sohrabi & Khanlari, 2007) و (Liu & Shih, 2005).

الگوریتم K-means

خوشه‌بندی فرآیندی است که مجموعه‌ای از موضوعات یا چیزهای فیزیکی یا غیرفیزیکی را در گروه‌های شبیه به هم گروه‌بندی می‌نماید (Han & Kamber, 2001). الگوریتم K-means یا روش فورگای (Forgy, 1965) از شناخته‌ترین الگوریتم‌های مورد استفاده در خوشه‌بندی است که کاربرد زیادی در آمار، تحلیل داده‌ها و مسائل تجاری دارد. این روش بر مبنای مقدار میانگین درون خوشه‌ها، آنها را به گونه‌ای افراز می‌نماید که عناصر داخل یک خوشه بیشترین نزدیکی را به میانگین همان خوشه داشته باشند (MacQueen, 1967). در این مقاله از این روش در ارتباط با خوشه‌بندی متغیرهای RFM استفاده می‌نمائیم. فرآیند محاسباتی K-means از قرار زیر است (Cheng & Chen, 2009):

۱- اقلام را در K خوشه اولیه قرار می‌دهیم.

⁶ - Recency, Frequency, Monetary

⁷ - Customer Life Time Value

۲- با انتصاب هر قلم به خوشه‌ای که کمترین فاصله مرکزی اقلیدسی را با آن دارد ادامه می‌دهیم. میانگین خوشه‌های با قلم اضافه شده و با قلم کسر شده را مجدداً حساب می‌کنیم.

۳- مرحله ۲ را آنقدر تکرار می‌کنیم تا انتصاب جدیدی قابل انجام نباشد. بجای افراز K خوشه در مرحله اول می‌توانیم مقادیر اولیه‌ای را بعنوان مقدار مرکزیت هر خوشه بطور دلخواه برای هر خوشه انتخاب کنیم.

اعداد فازی

اعداد فازی در اصل بیان و تعمیم اعداد اُردینال می‌باشند. یک عدد اُردینال مثل \tilde{a} می‌تواند با تابع عضویت زیر نشان داده شود:

$$\mu_{\tilde{a}}(x) = \begin{cases} 1 & ; \text{if } x = a \\ 0 & ; \text{if } x \neq a \end{cases} \quad (1)$$

بنابر این هر عدد حقیقی می‌تواند بعنوان یک عدد فازی بیان گردد که ساده‌ترین آنها اعداد فازی مثلثی می‌باشند (Jafari Samimi et al, 2010). تبدیل عدد فازی M را بر دامنه اعداد حقیقی R عدد فازی مثلثی می‌نامیم اگر تابع عضویت آن $\mu_{\tilde{a}}(x): R \rightarrow [0, 1]$ مساوی باشد با:

$$\mu_{\tilde{a}}(x) = \begin{cases} \frac{x}{m-l} - \frac{l}{m-l}, & x \in [l, m] \\ \frac{x}{m-u} - \frac{u}{m-u}, & x \in [m, u] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

اعداد فازی مثلثی می‌تواند با (l, m, u) بیان شود که پارامترهای l و m و u به ترتیب بیانگر کمترین مقدار محتمل، بیشترین مقدار متیقن و بیشترین مقدار محتمل می‌باشند که یک واقعه فازی را توصیف می‌کنند (Ertugrul & Karakasoglu, 2009). دو عمل مهم بر اعداد فازی که در این مقاله از آنها استفاده می‌کنیم از این قرار است که اگر دو عدد فازی مثلثی مثبت (l_1, m_1, u_1) و (l_2, m_2, u_2) داشته باشیم، سپس:

$$(l_1, m_1, u_1) \cdot (l_2, m_2, u_2) = (l_1 \cdot l_2, m_1 \cdot m_2, u_1 \cdot u_2) \quad (3)$$

$$(l_1, m_1, u_1)^{-1} \approx \left(\frac{1}{u_1}, \frac{1}{m_1}, \frac{1}{l_1} \right) \quad (4)$$

AHP فازی

روش تحلیلی سلسله مراتبی AHP از معروفترین روش‌های تصمیم‌گیری چندشاخصه است که توسط ساعتی در دهه ۱۹۷۰ ابداع گردید. شاخص‌ها می‌توانند کمی یا کیفی باشند. اساس روش تحلیلی سلسله مراتبی بر مقایسات زوجی نهفته است. در این روش تصمیم‌گیر با ساختن درخت سلسله مراتب تصمیم خود، شاخص‌ها و گزینه‌های تصمیم‌گیری را مشخص می‌نماید. سپس با انجام یک سری مقایسات زوجی وزن هر یک از فاکتورها را در راستای گزینه‌های رقیب مشخص می‌سازد (Saaty, 1980).

روش سنتی AHP به دلیل اینکه مقادیر دقیقی را برای نظر تصمیم‌گیر در مورد گزینه‌های مختلف در نظر می‌گیرد دچار اشکال می‌شود (Wang & Chen, 2007). و نتیجتاً نمی‌تواند در مقیاس‌های ناموزون قضاوت‌ها پاسخگو باشد و مقایسات زوجی را در زمان عدم دقت و نااطمینانی را صحیح انجام دهد (Deng, 1999). برای فائق

آمدن به این مشکل FAHP برای مسائل سلسله مراتبی ابداع شد. برای تصمیم گیر بسیار مورد اعتمادتر است اگر از قضاوت فاصله‌ای استفاده کند تا قضاوت نقطه‌ای (Kahraman et al, 2003).

در این مقاله از FAHP استفاده می‌کنیم که ابتدا توسط Chang (1996) مطرح شد. فرض کنید: $X = \{x_1, \dots, x_n\}$ مجموعه‌ای از ارقام (معیارها) باشد و $G = \{g_1, \dots, g_n\}$ مجموعه هدف باشد. در روش تحلیل گستره^۸ چانگ برای هر معیار، تحلیل گستره برای هر هدف یک به یک انجام می‌شود و بنابر این، m مقدار تحلیل گستره برای هر معیار بدست می‌آید:

$$M_{gi}^1, \dots, M_{gi}^m, \quad i = 1, \dots, n \quad (5)$$

که $M_{gi}^j (j = 1, 2, \dots, m)$ همگی اعداد فازی مثلثی هستند. مراحل به شکل زیر است:

مرحله ۱. مقدار گستره مصنوعی برای معیار i ام به شکل زیر تعریف می‌شود:

$$S_i = \sum_{j=1}^m M_{gi}^j \otimes g_i \left[\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j \right]^{-1} \quad (6)$$

علامت \otimes برای عمل ضرب گسترش یافته است. برای بدست آوردن $\sum_{j=1}^m M_{gi}^j$ عمل جمع فازی m مقدار تحلیل گستره برای یک ماتریس خاص به شکل زیر است:

$$\sum_{j=1}^m M_{gi}^j = (\sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m u_j) \quad (7)$$

برای بدست آوردن $[\sum_{j=1}^m M_{gi}^j]^{-1}$ ، عمل جمع فازی مقادیر $M_{gi}^j (j = 1, \dots, m)$ به صورت زیر است:

$$\sum_{i=1}^n \sum_{j=1}^m M_{gi}^j = (\sum_{i=1}^n l_i, \sum_{i=1}^n m_i, \sum_{i=1}^n u_i) \quad (8)$$

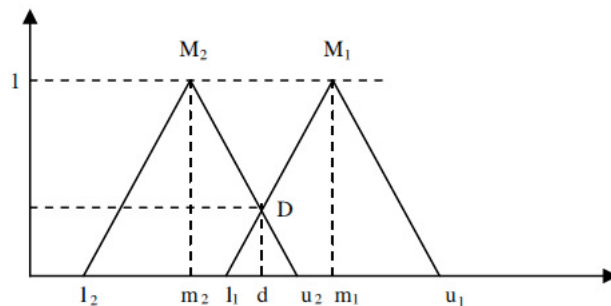
مرحله ۲. از جایی که $M_1 = (l_1, m_1, u_1)$ و $M_2 = (l_2, m_2, u_2)$ دو عدد فازی مثلثی هستند، درجه امکان^۹

$M_2 = (l_2, m_2, u_2) \geq M_1 = (l_1, m_1, u_1)$ را به شکل زیر تعریف می‌کنیم:

$$V(M_2 \geq M_1) = \begin{cases} 1 & \text{if } m_2 \geq m_1 \\ 0 & \text{if } l_1 \geq u_2 \\ \frac{l_1 - u_2}{(m_2 - u_1) - (m_2 - l_1)} & \text{otherwise} \end{cases} \quad (9)$$

برای مقایسه M_1 و M_2 به هر دو مقدار $V(M_2 \geq M_1)$ و $V(M_1 \geq M_2)$ احتیاج داریم. تصویر ۱ نشان می‌دهد که d

عرض بالاترین تقاطع در نقطه D بین μ_{M_1} و μ_{M_2} می‌باشد.



تصویر ۱. تقاطع بین دو عدد فازی $(M_1 \& M_2)$ و درجه امکان $M_1 \geq M_2$

⁸ Chang's Extent Analysis (EA)

⁹ Possibility degree

مرحله ۳. درجه امکان برای یک عدد فازی محدب که بزرگتر از k عدد فازی محدب M_i ($i = 1, \dots, k$) باشد می تواند به شکل زیر تعریف شود:

$$V(M \geq M_1, \dots, M_k) = V[(M \geq M_1), \dots, (M \geq M_k)] = \min V(M \geq M_i), i = 1, \dots, k \quad (10)$$

فرض کنید: $d(A_i) = \min V(S_i \geq S_k)$ for $k = 1, \dots, n; k \neq i$. بردار اوزان به شکل زیر خواهد بود:

$$w' = (d'(A_1), \dots, d'(A_n))^T \quad (11)$$

که A_i ($i = 1, \dots, n$)، n عضو آن هستند.

مرحله ۴. با نرمالیزه کردن، بردارهای اوزان بهنجار بدست می آید:

$$w = (d(A_1), \dots, d(A_n))^T \quad (12)$$

که W یک عدد نافازی است.

روش محاسبه

به منظور رتبه بندی اعتباری مشتریان حقوقی بانک توسعه صادرات در این مقاله، از مدل RFM استفاده شد. در رتبه بندی بر اساس RFM نگرشهای مالی چندان مطرح نیست و گرایش اصلی روش، به سمت مسائل کیفی است. سه فاکتور حجم مبادله، تعداد تکرار مبادله و تازگی مبادله در نظر گرفته شد که عموماً به صورت مستقیم با سوددهی مشتری ارتباطی ندارد. به همین دلیل در این تحقیق در کنار متغیرهای RFM، از داده های تراکنشی مشتری و رفتار مشتری در بازپرداخت وام نیز استفاده گردید. تعداد ۲۹۶ مشتری حقوقی بانک در سال ۱۳۸۸ به عنوان نمونه در نظر گرفته شد و ۵ متغیر مدل به صورت زیر محاسبه شدند:

متغیر تأخر (R) فاصله بین تازه ترین زمان تراکنش و زمان تحلیل را اندازه می گیرد. در واقع R آخرین باری را که در طی بازه زمانی مورد نظر، مشتری صادرات^۱ داشته است، را مشخص می کند. متغیر تکرر (F)، تکرار صادرات در دوره زمانی مورد نظر را اندازه می گیرد. ارزش پولی (M)، کل ارزش دلاری صادرات در طی بازه مفروض را مشخص می کند. داده های تراکنشی، مربوط به تعداد تراکنش های مشتری در دوره زمانی مورد نظر بوده و رفتار مشتری در باز پرداخت وام، مربوط به تعداد وام های معوق و یا سر رسید گذشته مشتری به بانک یا سایر بانک های کشور می باشد. متغیرهای مورد نظر پس از استخراج از بانک اطلاعاتی بانک توسعه صادرات، نرمالیزه گردیدند. همچنین وزن هر یک از متغیرها با استفاده از روش تحلیل سلسله مراتبی فازی محاسبه و نهایتاً ارزش هر مشتری برآورد گردید. در ادامه مشتریان بانک با استفاده از روش خوشه بندی K-means رتبه بندی شدند.

ابتدا به منظور برآورد وزن هر یک از متغیرها، پرسشنامه ای بر اساس روش تحلیل سلسله مراتبی تهیه و در بین ۳۰ نفر از کارشناسان بانکی توزیع و تکمیل گردید. در این پرسشنامه، متغیرها به صورت زوجی با هم مقایسه شدند. پس از تبدیل متغیرهای زبان شناختی به اعداد فازی مثلثی، میانگین داده های جمع آوری شده از سی نمونه محاسبه و داده های تلفیقی فازی بدست آمد که در جدول (۱) نشان داده شده اند. در این جدول R, F, M, TD, RB به ترتیب، رفتار بازپرداخت، تراکنش، پولی، تکرر و تأخر می باشند:

^۱ مشتریان بانک توسعه صادرات از منابع دریافتی جهت صادرات استفاده نموده اند.

جدول (۱) ماتریس اولیه مقایسه زوجی متغیرها پس از تلفیق ۳۰ داده

متغیرها	R	F	M	TD	RB
R	(1,1,1)	(1,2.33,3)	(3,3.67,5)	(0.33,4.11,7)	(0.14,4.05,7)
F	(0.33,0.55,1)	(1,1,1)	(0.33,1.44,3)	(0.14,1.78,5)	(0.2,1.18,3)
M	(0.2,0.28,0.33)	(0.33,1.44,3)	(1,1,1)	(0.14,1.11,3)	(0.14,1.76,5)
TD	(0.14,1.11,3)	(0.2,4.07,7)	(0.33,4.11,7)	(1,1,1)	(3,3,3)
RB	(0.14,2.45,7)	(0.33,2.78,5)	(0.2,4.73,7)	(0.33,0.33,0.33)	(1,1,1)

پس از تلفیق داده‌ها، بردارهای S_i به صورت زیر محاسبه گردید:

$$S_1 = (5.47, 15.16, 23) \otimes \left(\frac{1}{89.66}, \frac{1}{51.28}, \frac{1}{16.06} \right) = (0.06, 0.29, 1.43) \quad (13)$$

$$S_2 = (2, 5.95, 13) \otimes \left(\frac{1}{89.66}, \frac{1}{51.28}, \frac{1}{16.06} \right) = (0.02, 0.12, 0.81)$$

$$S_3 = (1.81, 5.59, 12.33) \otimes \left(\frac{1}{89.66}, \frac{1}{51.28}, \frac{1}{16.06} \right) = (0.02, 0.11, 0.77)$$

$$S_4 = (4.67, 13.29, 21) \otimes \left(\frac{1}{89.66}, \frac{1}{51.28}, \frac{1}{16.06} \right) = (0.05, 0.26, 1.31)$$

$$S_5 = (2, 11.29, 20.33) \otimes \left(\frac{1}{89.66}, \frac{1}{51.28}, \frac{1}{16.06} \right) = (0.02, 0.22, 1.26)$$

در گام بعدی درجه امکان (برای بردارهای S_i) نسبت به هم با استفاده از فرمول (۹) محاسبه گردید. در گام

بعدی مقادیر $d(I)$ ها به صورت زیر محاسبه گردید:

$$d'(I1) = \text{MIN}(S_1 \geq S_2, S_3, S_4, S_5) = \text{MIN}(1, 1, 1, 1) = 1 \quad (14)$$

$$d'(I2) = \text{MIN}(S_2 \geq S_1, S_3, S_4, S_5) = \text{MIN}(0.82, 1, 0.84, 0.89) = 0.82$$

$$d'(I3) = \text{MIN}(S_3 \geq S_1, S_2, S_4, S_5) = \text{MIN}(0.8, 0.99, 0.83, 0.87) = 0.8$$

$$d'(I4) = \text{MIN}(S_4 \geq S_1, S_2, S_3, S_5) = \text{MIN}(0.98, 1, 1, 1) = 0.98$$

$$d'(I5) = \text{MIN}(S_5 \geq S_1, S_2, S_3, S_4) = \text{MIN}(0.94, 1, 1, 0.97) = 0.94$$

و نهایتاً بردار نهائی به صورت زیر محاسبه گردید:

$$W' = (1, 0.82, 0.8, 0.98, 0.94)^T \quad (15)$$

$$W = (0.22, 0.18, 0.176, 0.216, 0.21) \quad (16)$$

لذا بر پایه روش FAHP اولویت بندی متغیرها به صورت زیر بدست آمد:

جدول (۲) ماتریس نهائی اولویت بندی متغیرهای مدل با استفاده از روش FAHP

وزن معیارها	معیارها
0.220	تأخر
0.216	رفتار بازپرداخت
0.210	داده‌های تراکنشی
0.180	تکرر
0.176	پولی

پس از مشخص شدن وزن شاخص‌ها، در گام بعدی مقادیر متغیرهای مدل با فرمول زیر نرمال شدند:

$$y_{ci} = \frac{x_{ci} - x_{Min}}{x_{Max} - x_{Min}} \quad (17)$$

x_{ci} مقدار اصلی متغیرها برای مشتری نام و x_{min} و x_{max} حداقل و حداکثر مقدار هر متغیر در بین کل مشتریان نمونه است. سپس با استفاده از الگوریتم K-means ۲۹۶ نمونه مورد مطالعه خوشه بندی شدند. ابتدا تعداد خوشه بهینه برآورد گردید که در جدول (۳) ذکر شده‌اند:

جدول (۳) نتایج الگوریتم خوشه بندی kmeans جهت تعیین تعداد خوشه‌ها

Kmeans	تعداد خوشه‌ها
۰/۲۲۰	۲
۰/۱۴۸	۳
۰/۱۲۹	۴
۰/۱۰۹	۵
۰/۱۳۱	۶

تعداد بهینه خوشه‌ها ۵ خوشه بدست آمد. سپس ارزش هر مشتری نیز بر اساس ۵ متغیر ارائه شده، به صورت

زیر محاسبه گردید:

$$V(C_i) = W_1 \times R(C_i) + W_2 \times F(C_i) + W_3 \times M(C_i) + W_4 \times TD(C_i) + W_5 \times RB(C_i) \quad (18)$$

که $R(C_i)$, $F(C_i)$, $M(C_i)$, $TD(C_i)$, $RB(C_i)$ به ترتیب، رفتار بازپرداخت، تراکنش، پولی، تکرر و تأخر را نشان می‌دهند. W_i وزن متغیرها را نشان می‌دهد که در مرحله قبل با استفاده از روش تحلیل سلسله مراتبی فازی برآورد گردیدند. کل مشتریان در ۵ خوشه رتبه بندی گردید. نتایج حاصله در جدول (۴) آورده شده است.

بر اساس نتایج بدست آمده، مشتریان حقوقی بانک در سال ۱۳۸۸ به ۵ خوشه اصلی تقسیم می‌شوند. خوشه سوم (با ارزش خوشه ۰/۹۵) و خوشه پنجم (با ارزش خوشه ۰/۷۶) جمعاً شامل ۱۶۲ شرکت (۵۴٪ مشتریان) می‌باشند، مشتریان خوب و بسیار خوب بانک را تشکیل می‌دهند. ریسک اعتباری این دو گروه (بالاخص خوشه سوم) در حداقل مقدار خود قرار دارد. خوشه دوم شامل ۹۶ شرکت (۳۲٪ مشتریان) و با ارزش خوشه ۰/۵۹ مشتریان متوسط بانک را به لحاظ اعتباری تشکیل می‌دهد. دو خوشه چهارم و اول نیز با ارزش خوشه ۰/۴۲ و ۰/۲۶ به ترتیب پر

ریسک‌ترین مشتریان بانک (۱۴٪ مشتریان) را تشکیل می‌دهند. پرداخت تسهیلات به این گروه‌ها به خصوص خوشه اول از احتمال نکول بالایی برخوردار است.

جدول شماره (۴) ارزش خوشه‌های تشکیل شده

شماره خوشه	میانگین R	میانگین F	میانگین M	میانگین RB	میانگین TD	ارزش هر خوشه	تعداد اعضای خوشه
3	0.82	0.91	1	1	1	0.95	34
5	0.57	0.83	0.78	0.72	0.88	0.76	128
2	0.38	0.69	0.64	0.62	0.64	0.59	96
4	0.24	0.54	0.56	0.52	0.31	0.42	20
1	0.19	0.47	0.34	0.31	0.19	0.26	18

نتیجه‌گیری

یکی از مشکلات تجزیه و تحلیل اطلاعات مشتریان بانکها، چند بعدی بودن اطلاعات آنان می‌باشد. مدل‌های رفتاری و اعتباری، با توجه به آنکه دو جنبه اصلی رفتار مشتریان را مد نظر قرار می‌دهند، از موفقترین مدل‌های کاربردی در ارزیابی مشتریان بانکها محسوب می‌شوند.

در این مقاله فرایندی تلفیقی از الگوهای اعتباری و رفتاری و الگوریتم K-means جهت دسته‌بندی اعتباری مشتریان بانکی ارائه شد. با توجه به اینکه در رتبه‌بندی بر اساس RFM، نگرشهای مالی چندان مطرح نیست، به همین دلیل در این تحقیق در کنار متغیرهای RFM، از داده‌های تراکنش مشتری و رفتار مشتری در بازپرداخت وام نیز استفاده گردید. نتایج روش پیشنهادی با توجه به سهولت استفاده و چند بعدی بودن اطلاعات، می‌تواند به بانکها و موسسات مالی کمک نماید تا مشتریان خود را به صورت هدفمند دسته‌بندی نمایند.

منابع

- غضنفری، مهدی و ملک محمدی، سمیرا؛ عزیززاده، سمیه؛ فتح الله، مهدی، بخش‌بندی مشتریان صادراتی میوه‌های خوراکی، فصلنامه پژوهشهای بازرگانی، شماره ۵۵ تابستان ۸۹، ۱۵۱-۱۸۱.
- رزمی، جعفر و ارش قنبری، ارائه مدلی نوین جهت محاسبه ارزش دوره عمر مشتری، فصلنامه مدیریت فناوری اطلاعات، دوره ۱، شماره ۲، بهار و تابستان ۱۳۸۸، ۳۵-۵۰.
- Chang, D. Y. (1996) Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research*, 95, 649–655.
- Chen, M. C., Huang, S. H. (2003) Credit scoring and rejected instances reassigning through evolutionary computation techniques. *Expert Systems with Applications*, 24, 433–441.
- Cheng, C.H., You-Shyang Chen. (2009) Classifying the segmentation of customer value via RFM model and RS theory. *Expert Systems with Applications* 36, 4176–4184.
- Deng, H. (1999) Multicriteria analysis with fuzzy pair-wise comparison. *International Journal of Approximate Reasoning*, 21, 215–231.
- Dirk Van den Poel, Bart Larivi, (2004), Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research*. 157, pp. 196–217.
- Ertugrul, I., N. Karakasoglu (2009), Performance evaluation of Turkish cement firms with fuzzy analytic hierarchy process and TOPSIS methods. *Expert Systems with Applications* 36, pp. 702–715.

- Forgy, E. (1965) Cluster analysis of multivariate data: Efficiency versus interpretability of classifications. *Biometrics*, 21, 768.
- Han, J., Kamber, M. (2001) *Data mining: Concepts and techniques*. San Francisco: Morgan Kaufmann Publishers.
- Hand, D. J. (1981) *Discrimination and classification*. New York: Wiley.
- Heckerman, D. (1996) Bayesian networks for knowledge discovery. *Advances in knowledge discovery and data mining*, pp. 273–305.
- Hsieh, Nan-Chen (2004) An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert Systems with Applications* 27, pp. 623–633.
- Hughes, A. M. (1994) *Strategic database marketing*. Chicago: Probus Publishing Company.
- Hyunseok Hwang, Taesoo Jung, Euiho Suh (2004). An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. *Expert Systems with Applications* 26, pp. 181–188.
- Jafari samimi, A., Bidabad, B., Mohammadi, R. (2010) Simulation of continuous qualitative variables in econometric models using fuzzy functions and numbers. *Australian Journal of Basic and Applied Sciences*, 4(10): 4780-4791.
- Johnson, R. A., Wichern, D. W. (1998) *Applied multivariate statistical analysis* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Jonkera, J-J., Piersmab, N. and Van den Poelc, D. (2004); Joint optimization of customer segmentation and marketing policy to maximize long-term profitability. *Expert Systems with Applications*, Vol. 27, pp.159–168.
- Kahraman, C., Cebeci, U., Ulukan, Z. (2003) Multi-criteria supplier selection using fuzzy AHP. *Logistics Information Management*, 16(6), 382–394.
- Lancher, R. C., Coats, P. K., Shanker, C. S., Fant, L. F. (1995) A neural network for classifying the financial health of a firm. *European Journal of Operational Research*, 85(1), 53–65.
- Liu, D., Ya-Yueh Shih (2005) Integrating AHP and data mining for product recommendation based on customer lifetime value, *Information and Management* 42(3), 387-400.
- MacQueen, J. B. (1967) Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley symposium on mathematical statistics and probability*, pp. 281–297. Berkeley: University of California Press.
- Morrison, D. F. (1990) *Multivariate statistical methods*. New York, NY: McGraw-Hill.
- Newell, F. (1997) *The new rles of marketing: how to use one-to-one relationship marketing to be the leader in your industry*. New York: McGraw-Hills Companies Inc.
- Saaty, T. L. (1980) *The analytic hierarchy process*. New York: McGraw- Hill.
- Setiono, R., Thong, J. Y. L., Yap, C. S. (1998) Symbolic rule extraction from neural networks, an application to identifying organizations adopting IT. *Information and Management*, 34(2), 91–101.
- Sohrabi, B. Amir Khanlari, (2007), Customer Lifetime Value (CLV) measurement based on RFM Model. *Iranian Accounting and Auditing Review*, Spring, Vol. 14 No. 47, pp. 7- 20.
- Thomas, L. C. (2000) A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16, 149–172.
- Wang, CH,. (2010) Apply robust segmentation to the service industry using kernel induced fuzzy clustering techniques. *Expert Systems with Applications*, 37: 8395-8400.
- Wang, T. C., Chen, Y. H. (2007) Applying consistent fuzzy preference relations to partnership selection. *Omega, the International Journal of Management Science*, 35, 384–388.
- Wu, J., Lin, Z. (2005) Research on customer segmentation model by clustering. *ACM International Conference Proceeding Series*, p. 113.
- Yeh, C., Yang, K. and Ting, T. (2009) Knowledge discovery on RFM model using Bernoulli sequence. *Expert Systems with Applications*, Vol. 36, pp. 5866–5871.