

A multi-objective model of service assignment to bank customers by data mining and optimization via simulation

(Document Type: Research Paper)

Seyed Mohammad Ali Khatami Firouzabadi*

Department of Industrial Management, Faculty of Management and Accounting,
Allameh Tabataba'i University, Tehran, Iran, smakhf@hotmail.com

Mohammad Taghi Taghavifard

Department of Industrial Management, Faculty of Management and Accounting,
Allameh Tabataba'i University, Tehran, Iran, dr.taghavifard@gmail.com

Khalil Sajjadi

Department of Industrial Management, Faculty of Management and Accounting,
Allameh Tabataba'i University, Tehran, Iran, khalil_sajjadi2006@yahoo.com

Jahanyar BamdadSoufi

Department of Industrial Management, Faculty of Management and Accounting,
Allameh Tabataba'i University, Tehran, Iran, bamdadsoofi@yahoo.com

Purpose: The main purpose of this paper is to propose a multi-objective model for assigning service/product to clustered customers. The main practical objectives of this model from the perspective of the bank are reduced cost and risk and increased customer satisfaction.

Design/methodology/approach: In this paper, five indicators of recency, frequency, monetary, loan and deferred have been identified and customers have been clustered, accordingly using K-means approach. Then, a three-objective mathematical model has been designed to assign optimal service/product as response to customer. Finally the model has been solved by simulation based optimization.

Findings: In the case study, all information about five characteristics of customers was extracted from the database, 31953 customers were placed in seven clusters and the validity of these clusters was measured. A three-objective mathematical model was designed based on the characteristics of 13 types of bank products/services. Then, the simulation modeling solutions were improved using the simulated annealing algorithm. In this study, Weka and R-Studio, Arena and Longo were used for data mining, simulation and optimization, respectively.

Research limitations/implications: The limitations of this study include inability of simulation instruments for drawing, solving all probable states (more scenarios) and solving the model for those states. It is recommended to develop the mathematical model with respect to customer, so that after problem solving, the bank would be able to make decision on providing services and products to its customers. Simultaneously, the objective functions would be fitted within their most reasonable states and ultimately, using a model, the parameters related to each product can be set for the new customer referring to the bank.

* Corresponding author

Copyright © 2020, University of Isfahan. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits others to download this work and share it with others as long as they credit it, but they cannot change it in any way or use it commercially.

Practical implications: Products/services were assigned according to customer needs in a way that cost and risk were reduced and the utility of assignment was increased through the proposed model and simulating the behavior of each cluster of customers.

Social implications: Paradigm shift in the banking industry is changing from e-banking to digital banking. In digital banking, assigning/customizing products/services, regarding the needs of customers, is very difficult. The banking industry is not well equipped to respond to the digital banking expectations of most consumers. One of the most important challenges of banks is recognizing customers, clustering and assigning a service/product to each of the different clusters. The main policy in the banking industry is to increase customer satisfaction and reduce cost and risk in sales service. Therefore, each customer should have a dedicated service/product.

Originality/value: In this paper, authors attempted to use one of the clustering approaches in multi-objective programming. In addition, they proposed an approach for assigning product/service to customer by simulating and analyzing the behavior of each customer cluster.

Keywords: Multi objective assignment model, Bank customers, Clustering, Optimization via simulation

مدیریت تولید و عملیات، دوره ۱۰، پیاپی ۱۹، شماره ۲، پاییز زمستان ۱۳۹۸

دریافت: ۱۳۹۶/۰۷/۲۰ پذیرش: ۱۳۹۸/۰۷/۱۵

صص: ۱۶۱-۱۸۰ (نوع مقاله: پژوهشی)

ارائه مدل بهینه‌سازی چندهدفه تخصیص خدمت به مشتریان بانک به کمک داده کاوی و شبیه‌سازی

سید محمدعلی خاتمی فیروزآبادی^{۱*}، محمدتقی تقوی فرد^۲، سید خلیل‌الله سجادی^۳، جهانیار بامدادصوفی^۴

۱- دانشیار گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران، smakhf@hotmail.com

۲- دانشیار گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران،

dr.taghavifard@gmail.com

۳- دانشجوی دکتری گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران،

khalil_sajjadi2006@yahoo.com

۴- استادیار گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی، تهران، ایران،

bamdadsoofi@yahoo.com

چکیده: امروزه شناخت مشتریان، خوشه‌بندی و تخصیص خدمت یا محصول به هرکدام از خوشه‌های مختلف یکی از مهم‌ترین مسائل بانک‌ها محسوب می‌شود. در این پژوهش اطلاعات ۳۱,۹۵۳ مشتری شامل پنج ویژگی، آخرین زمان مراجعه، تعداد تراکنش، مبلغ سپرده‌گذاری، مبلغ وام و مانده معوقات از پایگاه داده بانک استخراج شده است. سپس به کمک الگوریتم کا میانگین مشتریان در ۷ خوشه جایگذاری شده است. هدف اصلی این پژوهش تخصیص ۹ نوع وام و ۴ نوع سپرده به هر خوشه از مشتریان براساس یک مدل ریاضی سه‌هدفه برای افزایش میزان رضایت مشتریان، کاهش هزینه‌ها و ریسک تخصیص خدمات و محصولات است. برای حل این مدل جواب‌های موجه اولیه در قالب سناریوهای مختلف از طریق شبیه‌سازی به دست آمده است. سپس به کمک الگوریتم تبرید جواب نزدیک به بهینه مشخص شده است. در این پژوهش از نرم‌افزارهای وکا و آر برای داده‌کاوی، ارنای برای شبیه‌سازی و لینگو برای بهینه‌سازی استفاده شده است.

واژه‌های کلیدی: مدل تخصیص چندهدفه، مشتری، خوشه‌بندی، بهینه‌سازی، شبیه‌سازی

مقدمه

امروزه با گسترش بنگاه‌های مالی و اقتصادی، همچنین افزایش حجم، تنوع فعالیت‌ها و رقابت شدید در جذب مشتریان، نیاز به حفظ رضایت آنها بیش از پیش احساس می‌شود. در پژوهشی با توسعه یک مدل ریاضی میزان رضایت مشتریان بانک‌ها در استفاده از خدمات بانکی پیش‌بینی شده است (آبیدان^۱، ۲۰۱۷). یکی از این بنگاه‌های مالی بانک‌ها و مؤسسات اعتباری هستند که در حال حاضر نارضایتی گسترده مشتریان از نحوه ارائه خدمات و نوع محصولات ارائه شده به آنها مشاهده می‌شود؛ این امر بیشتر به لحاظ منطبق نبودن خدمات و محصولات بانکی با نیاز و توقعات مشتریان به وجود آمده است (اکبری اصل و بشلی، ۲۰۱۴). از سوی دیگر تعامل مشتریان با بانک‌ها باعث بروز وقایعی می‌شود که اطلاع از آنها، تجزیه و تحلیل و طبقه‌بندی آنها بسیار حائز اهمیت است. تقسیم‌بندی مشتریان به لحاظ رفتار بوجود آمده در هر بنگاه اقتصادی عامل اصلی در مدیریت ارتباط با مشتریان و بازاریابی هدفمند به شمار می‌رود. تاکنون روش‌های مختلفی برای تخصیص خدمت به مشتری ارائه شده است که بخش درخور توجهی از آنها به صورت تجربی است؛ به نحوی که بیشتر بانک‌ها با در نظر گرفتن تنها یک یا دو پارامتر مانند مانده میانگین و گردش حساب اقدام به بخش‌بندی مشتریان خود کرده‌اند و آنها را در گروه‌های پلاتینی، طلایی، نقره‌ای و غیره جای می‌دهند. برخلاف این مشکلات و پیچیدگی‌ها، مطالعات نشان می‌دهد بانک‌هایی که از بخش‌بندی مشتریان به درستی استفاده کرده‌اند ۴۰٪ درآمد بیشتری از سایر بانک‌ها داشته‌اند و درآمد مشتریان فعال این بانک‌ها ۲۰٪ از مشتریان فعال سایر بانک‌ها بیشتر است (کومار و همکاران^۲، ۲۰۱۲). شناسایی حساب‌های کلیدی^۳ افتتاح شده توسط مشتریان سودآور، می‌تواند یکی از مزایای بخش‌بندی صحیح و اصولی مشتریان در بانک باشد. فیشر برای نخستین بار سال ۱۹۶۳ ایده متمرکز کردن گروه‌ها را مطرح کرد. سپس دیویدراند در سال ۱۹۴۱ با تفکیک مشتریان به دو گروه خوب و بد به آنها وام اعطا کرد. بعد از آن در سال ۱۹۶۰ کارت‌های اعتباری برای هر گروه از مشتریان وارد بازار شد. در سال ۱۹۸۰ برای نخستین بار در آمریکا از اعتبارسنجی هر گروه از مشتریان استفاده شد. در سال ۱۹۹۰ معیار اصلی ارائه خدمت به مشتریان منوط به اعتبارسنجی مشتریان شد (توماس^۴، ۲۰۰۱). سپس در دهه ۹۰ میلادی کاتلر برای نخستین بار مباحث مربوط به مشتری‌مداری و مدیریت ارتباط با مشتری را مطرح کرد (کاتلر^۵، ۱۹۹۴). در امتداد این مباحث با ورود سیستم‌های کامپیوتری به بانک‌ها و بازارهای مالی، اهمیت استفاده از پایگاه‌های داده و به‌کارگیری آن در برنامه‌ریزی، گزارش‌گیری و شناسایی رفتار مشتریان روزبه‌روز بیشتر شد (آناناسوپلاس^۶ و همکاران، ۲۰۰۰). در پژوهشی با تحلیل رفتار مشتریان از طریق تراکنش‌های تلفن همراه، از وقوع جرم جلوگیری شده است (چن^۷ و همکاران، ۲۰۱۷). واضح است که حفظ یا افزایش رضایت مندی مشتریان ارتباط مستقیمی با افزایش کیفیت خدمات و یا افزایش پارامترهای مدنظر مشتری در بهره‌مندی از محصولات بانک دارد و اگر بانک یا هر بنگاه مالی به دنبال این هدف باشد، متحمل هزینه خواهد شد. در نهایت این نوع هزینه‌ها باعث کاهش سود بنگاه می‌شود. هدف اصلی این مقاله ارائه محصولات و خدمات بانکی از طریق خوشه‌بندی مناسب مشتریان براساس تحلیل رفتاری پنج ویژگی آنها با در نظر گرفتن هدف افزایش میزان رضایت مندی مشتری، کاهش هزینه‌ها و ریسک بانک در حالتی است که بیشترین حد مطلوبیت هریک از اهداف حاصل شود. به عبارت دیگر این مقاله با تخصیص بهینه خدمات و محصولات به خوشه‌های مشتریان هدف، به‌طور هم‌زمان افزایش رضایت مندی مشتریان در بهره‌مندی از خدمات و محصولات بانک و کاهش هزینه‌ها و ریسک ناشی از تخصیص مناسب خدمات و محصولات به مشتریان را مدنظر دارد.

در این بخش سعی شده است با مطالعه گسترده در حوزه ادبیات موضوع، پیشینه جامعی گردآوری و توسعه و کاربرد مدل‌های موجود در این حوزه ارائه شود. از آنجایی که شناسایی رفتار مشتریان و تخصیص خدمت به آنان تا حدی ریشه در مدیریت ارتباط با آنها دارد، پیشینه مدیریت ارتباط با مشتری همواره رویکردی مهم در کسب و کار است و به دهه ۱۹۹۰ مربوط است (انگای^۸، ۲۰۰۵). در تعاریف، امروزه تأکید زیادی بر اهمیت شناخت رفتار مشتری (فرایندی جامع و استراتژیک برای حداکثرسازی ارزش آنان برای سازمان‌ها) شده است (میشار^۹، ۲۰۰۹). از آنجاکه سازمان‌ها به مشتری‌محوری توجه زیادی دارند، تعداد مشتریان به صورت فزاینده‌ای رشد کرده است؛ اما به دنبال آن سازمان‌ها قادر به خدمت‌دهی به صورت یکسان به همه مشتریان نبوده‌اند؛ بنابراین لازم است مشتریان ارزش‌گذاری شوند و با توجه به اهمیت و ارزش هر کدام، نوع خدمات متناسب آنها ارائه شود تا تشویقی برای ارتباط مستمر مشتری با سازمان باشد؛ در نتیجه سازمان مشتریان هدف خود را به خوبی شناسایی کند و ارتباطی درخور آنان داشته باشد؛ بنابراین مباحثی با عنوان ارزش دوره عمر مشتری، حقوق مشتری، سودآوری مشتری، طبقه‌بندی مشتری و بخش‌بندی در مطالعات متعدد بررسی شده است (وانگ^{۱۰} و همکاران، ۲۰۰۶). در واقع حرکت به سمت بازاریابی مشتری‌محور همراه با افزایش دسترسی به اطلاعات و تراکنش‌های مشتریان موجب علاقه زیاد به اندازه‌گیری ارزش دوره عمر مشتری و طبقه‌بندی آنان شده است (صفری کهره، ۲۰۱۲). در پژوهشی ضمن معرفی روش‌های محاسبه ارزش طول عمر مشتریان، از مدل پایه RFM استفاده شده است و با اضافه کردن اندازه ریسک مشتریان و شاخص‌های مالی دیگر با رویکردی غیرپارامتریک در فضای پژوهش امتیازدهی و بخش‌بندی مشتریان انجام شده است (سینگ^{۱۱}، ۲۰۱۶). در پژوهش دیگری کاربرد روش RFM برای شناسایی رفتار مشتریان سیستم‌های ارتباطی به کار می‌رود که در آستانه ورشکستگی هستند (زابکوسکی^{۱۲}، ۲۰۱۶).

روش‌های زیادی برای تخصیص خدمت به مشتری مبتنی بر ارزش دوره عمر مشتری (CLV^{۱۳}) و طبقه‌بندی مشتریان مطرح شده است. اساس برخی از آنها بر پایه مدل‌های معروفی نظیر ارزش فعلی خالص (NPV^{۱۴})، روش سهم کیف پول (SOW^{۱۵})، روش زنجیره مارکوف^{۱۶}، روش ارزش گذشته مشتری (PCV^{۱۷})، روش بازگشت سرمایه (ROI^{۱۸}) و روش RFM^{۱۹} است. در میان روش‌های مذکور مدل RFM یکی از پرکاربردترین و متداول‌ترین روش‌هایی است که در تعیین ارزش مشتری سه معیار را بررسی کرده‌اند (باتل^{۲۰}، ۲۰۰۴). در پژوهش‌های داخلی و خارجی با به‌کارگیری این مدل ارزش، طبقه‌بندی و بخش‌بندی مشتریان محاسبه شده است و نتایج درخور توجهی ارائه شده است؛ به طوری که در مقاله‌ای با به‌کارگیری این مدل، مشتریان خاص خود را در صنعت ارزشیابی و مدیریت کرده‌اند (کینینگهام^{۲۱} و همکاران، ۲۰۰۶). در جایی دیگر با به‌کارگیری این مدل و روش آماری^{۲۲} مشتریان دسته‌بندی آنان تحلیل شده است (وو^{۲۳} و همکاران، ۲۰۰۹). در پایان‌نامه‌ای با عنوان "الگوی برای تعیین ارزش چرخه عمر مشتریان" با استفاده از اطلاعات تراکنشی مربوط به ۵۰۰۰ مشتری حقیقی و حقوقی بانک ملت مدلی برای سنجش و اندازه‌گیری ارزش طول عمر مشتری ارائه شده است. هر گروه از مشتریان حقیقی و حقوقی به صورت جداگانه به ۹ بخش تقسیم شده‌اند. سپس محاسبه ارزش طول عمر مشتری برای هر بخش از مشتریان انجام شده است (سهرابی، خانلری و آجرلو، ۲۰۱۱). پایان‌نامه‌ای دیگر در راستای بررسی رفتار مشتریان در میزان استقبال از کانال‌های ارتباطی نوین بانکداری الکترونیک در بانک صادرات ارائه شده است. در این پژوهش رفتار

مشتری در استفاده از این ابزارها تحلیل شده است و با استفاده از روش‌های داده‌کاوی، مشتریان بخش‌بندی شده‌اند (تقوی فرد و خواجهوند، ۲۰۱۳). در مقاله‌ای دیگر با ارائه الگویی جامع ارزش چرخه عمر مشتری محاسبه شده است. هم‌چنین از عوامل مؤثر بر جریان نقدی حاصل از مشتری در دوره معین، نرخ نگهداری و تنزیل و نرخ رشد جریان نقدی مشتری استفاده شده است. در نهایت با استفاده از این مدل و به‌کارگیری روش‌های داده‌کاوی، بخش‌بندی حدود ۳۰۰۰ مشتری در بانک صادرات شعب تهران انجام شده است. در نتیجه برای چهار گروه مشتریان تعیین شده برنامه‌های بازاریابی و توسعه محصول تدوین شده است (تقوی فرد و همکاران، ۲۰۱۱). در مقاله‌ای دیگر با استفاده از مدل RFM و شبکه عصبی مشتریان اعتباری بانک بخش‌بندی و رتبه‌بندی شده‌اند. در نتیجه مشتریان به ۱۰ بخش تقسیم شده‌اند و براساس امتیاز بخش‌ها به مشتریان تسهیلات اعطا شده است (افسر، امیر و همکاران، ۲۰۱۴). البرزی و همکاران (۲۰۱۰) در مقاله خود با استفاده از روش Simple K-means مشتریان را به دو خوشه تقسیم‌بندی کرده‌اند. سپس با استفاده از الگوریتم ژنتیک ویژگی‌های اعتبارسنجی مهم در مجموعه داده انتخاب شده است و بهترین درخت تصمیم در هر خوشه مبتنی بر معیارهای بهینگی در هر خوشه ساخته شده است (البرزی و همکاران، ۲۰۱۰).

در پژوهشی دیگر انتظارات کاربران بانکداری تلفن همراه در قالب ۴ عامل خدمات اصلی، خدمات با ارزش افزوده، شرایط کاربرپسند و احساس رضایت در هنگام استفاده از خدمات با رویکردی کیفی و کمی شناسایی شده است. پس از مشخص کردن تعداد بهینه خوشه‌ها با روش وارد، از فن داده‌کاوی کامیاب‌ترین برای خوشه‌بندی استفاده شده است (عزیزی و بلاغی، ۲۰۱۶).

در مقاله دیگری با عنوان "بخش‌بندی کاربران بانکداری تلفن همراه بر مبنای انتظارات با رویکرد خوشه‌بندی"، رویکردی جامع برای بخش‌بندی بازار و طبقه‌بندی مشتریان با استفاده از روش‌های داده‌کاوی و برنامه‌ریزی خطی معرفی شده است. در این پژوهش الگوریتم یکپارچه‌ای برای بخش‌بندی بازار و طبقه‌بندی مشتریان ارائه شده است که خوشه‌بندی تجمیعی مبنای مقایسه عملکرد روش‌های مختلف طبقه‌بندی به‌ویژه برنامه‌ریزی خطی قرار می‌گیرد. این مقایسه نشان می‌دهد روش برنامه‌ریزی خطی عملکرد بهتری از نظر درصد بیشتر انتساب مشتریان به خوشه‌ها دارد. هم‌چنین برای مواجهه با ناکافی بودن داده‌های لازم در پایگاه داده، روش دلفی فازی پیشنهاد شده است (ایزدی و همکاران، ۲۰۱۶).

در مقاله دیگری ناظمی و همکارانش (۲۰۱۲) با استفاده از روش‌های داده‌کاوی مدلی دومرحله‌ای برای شناسایی ویژگی‌های گروه‌های مختلف مشتریان بانک توسعه داده‌اند و مشتریان مختلف از نظر میزان سودآوری برای بانک شناسایی شده‌اند (ناظمی و همکاران، ۲۰۱۲). در مقاله دیگری تقوی فرد و همکاران (۲۰۱۳) با استفاده از تحلیل RFM و به‌کارگیری الگوریتم two step خوشه‌بندی مشتریان را در چهار خوشه انجام داده‌اند و هر کدام را تحلیل کرده‌اند (تقوی فرد و همکاران، ۲۰۱۳). در اثر دیگری برادران و فرخی (۲۰۱۳) با استفاده از مدل توسعه‌یافته FRM و به‌کارگیری گام‌های CRISP-DM با به‌کارگیری الگوریتم دومرحله‌ای مشتریان یکی از بانک‌های خصوصی را بخش‌بندی کرده‌اند (برادران و فرخی، ۲۰۱۴). در اثری دیگر قربان‌پور و همکاران (۲۰۱۵) با تلفیق الگوریتم‌های ژنتیک و C-Means در محیط فازی مشتریان بانک رفاه را خوشه‌بندی کرده‌اند (قربان‌پور و همکاران، ۲۰۱۵).

مقایسه این پژوهش با سایر پژوهش‌های مطالعه‌شده نشان می‌دهد بیشتر مطالعات قبلی پس از انجام خوشه‌بندی مشتریان، با انجام تحلیل‌های موردی و کوتاه به اتمام رسیده‌اند؛ در حالی که مسئله مدنظر در این پژوهش با انجام فرایند خوشه‌بندی شروع و نتایج آن به‌عنوان ورودی به مدل برای به‌کارگیری در مراحل بعدی استفاده شده است و نیاز اصلی پژوهش را به‌صورت جامع پوشش می‌دهد. همچنین این پژوهش با در نظر گرفتن فرایندی جامع، مدلی را طراحی کرده است که اهداف اساسی بانک در عرصه خدمات و محصولات به مشتری را پوشش می‌دهد. در نهایت با استفاده از مفاهیم موجود در روش‌های بهینه‌سازی از طریق شبیه‌سازی مسئله حل شده است. این فرایند جامع در مطالعات پیشین به‌صورت متمرکز مطرح نشده است. در واقع در بیشتر مطالعات پیشین فقط جنبه علمی پژوهش پررنگ‌تر بوده است و بیان مطالعه موردی در پژوهش‌های پیشین جنبه تأیید روش و مدل حل مسئله است؛ در حالی که در این پژوهش، در مرحله نخست از مدل مطرح‌شده برای خوشه‌بندی مشتریان استفاده شده است. در ادامه از نتایج مدل چندهدفه تخصیص محصول به مشتری برای توسعه برنامه نرم‌افزاری هوش تجاری برای تسهیل در اتخاذ تصمیم برای ارائه خدمت و محصول بانک به مشتری از سوی مدیران و روسای شعب استفاده شده است. این نرم‌افزار که به‌وسیله یک شرکت نرم‌افزاری تهیه شده است در یکی از بانک‌های خصوصی کشور پیاده‌سازی و استفاده شده است.

مبانی نظری

خوشه‌بندی

خوشه‌بندی، گروه‌بندی یک سری موجودیت در دسته‌های مختلف است؛ این دسته‌ها نشان‌دهنده مفهوم یا معنی خاصی هستند؛ به عبارت ساده‌تر به یکدیگر شبیه باشند. از دیدگاه علمی تر خوشه‌بندی یعنی گروه‌بندی داده‌ها به k خوشه مختلف به طوری که داده‌هایی که در یک خوشه قرار می‌گیرند به یکدیگر شبیه باشند و داده‌های خوشه‌های مختلف با یکدیگر تفاوت داشته باشند. مدل‌های خوشه‌بندی به سه دسته کلی، سلسله مراتبی، مبتنی بر شبکه و مبتنی بر بخش‌بندی (تابع هدف) تقسیم می‌شوند. هریک از این دسته مدل‌ها ویژگی‌های خاص خود را دارند و ممکن است با توجه به کاربرد، متفاوت عمل کنند (مومنی، ۲۰۱۲).

روش خوشه‌بندی کا میانگین

این الگوریتم یکی از آسان‌ترین الگوریتم‌هایی است که برای حل مسائل مشهور دسته‌بندی طراحی شده است و مک کوئین در سال ۱۹۶۷ نخستین بار آن را در دانشگاه کالیفرنیا ارائه داده است. در این روش از معیار حداقل واریانس (یکی از معمول‌ترین روش‌ها برای خوشه‌بندی داده‌ها) استفاده شده است. برای این حداقل‌سازی، الگوهای مختلفی وجود دارد که یکی از عمومی‌ترین آنها روش خوشه‌بندی k -mean است. اگر مجموعه مشاهدات شامل X_1, X_2, \dots, X_n باشند و هر مشاهده یک بردار حقیقی d بعدی باشد، هدف خوشه‌بندی k -means بخش‌بندی این n مشاهده به k زیرمجموعه به‌گونه‌ای است که $S = \{S_1, S_2, \dots, S_k\}$ $k \leq n$ و مجموع مربعات میانگین هر خوشه حداقل شود. در آن میانگین نقاط در S_i است (شهرابی و ذوالقدرشجاعی، ۲۰۱۱).

روش پژوهش

روش انجام پژوهش از نقطه نظر هدف، توسعه‌ای-کاربردی و از نظر اجرایی تحلیلی-توصیفی است. جامعه هدف این پژوهش کلیه داده‌های ثبت شده برای مشتریان حقیقی بانک است. سؤال اصلی این پژوهش عبارت است از "مدل چندهدفه تخصیص خدمات و محصولات خاص به مشتریان حقیقی و حقوقی بانک چگونه است؟" به دنبال آن سؤالات فرعی عبارتند از ۱-خوشه‌بندی مشتریان با استفاده از داده‌کاوی براساس اهداف مدنظر بانک چگونه است؟ ۲- چگونه به کمک بهینه‌سازی از طریق شبیه‌سازی اهداف مدنظر بانک و مشتری تأمین می‌شود؟ ۳- سناریوهای پیشنهادی تخصیص خدمات به مشتریان کدام است؟ گردآوری داده‌ها و اطلاعات در این پژوهش از طریق نرم‌افزار SQLServer روی پایگاه داده بانک انجام شده است. همچنین از برنامه‌ها و نرم‌افزارهای مخصوص در حوزه داده‌کاوی از قبیل نرم افزار ۳,۸,۰ weka و R-Studio استفاده شده است. جامعه آماری این پژوهش تمامی ۸۵۰ هزار مشتری حقیقی بانک است که پس از پاک‌سازی داده‌ها اطلاعات مربوط به ۳۱,۹۵۳ مشتری حقیقی برای استفاده به کار گرفته شد. با توجه به تنوع زیاد ویژگی‌های ثبت شده از مشتریان و کاربردی بودن برخی از این ویژگی‌ها برای تحلیل رفتار ایشان، در این پژوهش مدل RFM توسعه یافته است. در واقع از طریق اضافه کردن دو ویژگی تسهیلات دریافتی و معوقات مشتریان به مدل RFM، ویژگی‌های مدنظر برای انجام فرایند خوشه‌بندی محاسبه شده است.

معرفی شاخص‌ها

از آنجاکه برای هریک از پنج شاخص مدنظر انواع مختلفی وجود دارد، سعی شده است تا با استفاده از روش مجموع ساده وزین (SAW) و در ادامه با به کارگیری روش آنتروپی برای تعیین اوزان، مقدار شاخص مربوطه به صورت ذیل محاسبه شود. نکته حائز اهمیت آنکه تمامی شاخص‌ها از عملکرد مالی مشتریان استخراج شده است. ۱- تاریخ آخرین مراجعه (R) ^{۲۴} (تراکنش) مشتری: از آنجاکه برای بانک انجام تراکنش‌های مالی از طریق کانال‌های مدرن دارای مطلوبیت بیشتری نسبت به مراجعات حضوری به شعبه است، در این مدل فاصله زمانی آخرین تراکنش مشتری از انتهای بازه زمانی مدنظر با نماد R و برحسب روز نشان داده می‌شود. همچنین از آنجاکه در حالت کلی تراکنش‌های مشتری تنها روی دو بستر (کانال) مدرن (خدمات غیرحضوری) و شعبه‌ای (مراجعات حضوری) رخ می‌دهد، برای آخرین تراکنش روی کانال‌های مدرن وزن α_{Rm} و با نماد R_m و برای مراجعه حضوری وزن α_{Rb} و با نماد R_b نمایش داده می‌شود؛ بنابراین مقدار R مشتری نام برابر است با رابطه ۱ که به صورت زیر نشان داده می‌شود.

$$R_i = (\alpha_{Rm} \times R_{mi}) + (\alpha_{Rb} \times R_{bi}) \quad \text{رابطه ۱}$$

۲- تعداد تراکنش‌ها در طول بازه زمانی (F) ^{۲۵}: مجموع تعداد کل تراکنش‌های مشتری در طول بازه زمانی مدنظر به دست می‌آید و با نماد F، برحسب تعداد بیان می‌شود.

۳- مجموع مانده سپرده‌ها (M) ^{۲۶}: از آنجاکه تمامی منابع در بانک به دو گروه "قرض الحسنه" و "سرمایه‌گذاری" تقسیم می‌شوند، مجموع کل مانده سپرده‌های مشتری نام با نماد M_i بیان می‌شود. مانده سپرده‌های جاری مشتری نام (منابع قرض الحسنه) با ضریب α_{Ml} و با نماد M_l و مجموع مانده سایر سپرده‌های مشتری نام با هزینه زیاد با ضریب α_{Mh} و با نماد M_h در نظر گرفته می‌شود. بنابراین مقدار M مشتری نام به صورت رابطه ۲ محاسبه می‌گردد.

$$M_i = (\alpha_{Mi} \times M_{li}) + (\alpha_{Mh} \times M_{hi}) \quad \text{رابطه ۲}$$

۴- مجموع مانده کل تسهیلات جاری (L)^{۲۷}: از آنجا که تسهیلات فعال (تسویه‌نشده) مشتری از نظر اهمیت برای بانک متفاوت است، مجموع آن در این پژوهش با نماد L نمایش داده می‌شود. در بانک تمامی تسهیلات و تعهدات در قالب سه بخش قرار می‌گیرد که هر کدام از آنها در ادبیات بانکی معنا و مفهوم خاص خود دارند و برای هر بخش وزن خاصی در نظر گرفته می‌شود. این اوزان به شرح رابطه ۳ است.

$$\begin{aligned} L_{mi} &: \text{مجموع تسهیلات مشارکتی مشتری نام با وزن } \alpha_{Lm} \\ L_{zi} &: \text{مجموع تعهدات (اعتباراسنادی و ضمانت‌نامه) مشتری نام با وزن } \alpha_{Lz} \\ L_{ni} &: \text{مجموع تسهیلات مبادله‌ای مشتری نام با وزن } \alpha_{Ln} \end{aligned}$$

$$L_i = (\alpha_{Lm} \times L_{mi}) + (\alpha_{Lz} \times L_{zi}) + (\alpha_{Ln} \times L_{ni}) \quad \text{رابطه ۳}$$

۵- مجموع کل معوقات (D)^{۲۸}: مجموع کل معوقات مشتری در پایان دوره در حال بررسی با نماد D نمایش داده و از سه بخش تشکیل می‌شود. مجموع کل معوقات در ادبیات بانکی معنا و مفهوم خاص خود را دارند و برای هر بخش وزن خاصی در نظر گرفته شده است که به شرح رابطه ۴ است.

$$\begin{aligned} D_{oi} &: \text{مجموع مانده مطالبات مشکوک الوصول بانک از مشتری نام با وزن } \alpha_{Do} \\ D_{di} &: \text{مجموع مانده مطالبات معوق بانک از مشتری نام با وزن } \alpha_{Dm} \\ D_{mi} &: \text{مجموع مانده مطالبات سررسید گذشته بانک از مشتری نام با وزن } \alpha_{Dm} \end{aligned}$$

$$D_i = (\alpha_{Do} \times D_{oi}) + \alpha_{Dm} \times (D_{di} + D_{mi}) \quad \text{رابطه ۴}$$

آماده‌سازی داده‌ها

پس از تعیین نحوه محاسبه متغیرها باید آماده‌سازی و پیش‌پردازش داده‌ها برای هریک از متغیرها انجام شود. آماده‌سازی و پیش‌پردازش داده‌ها طی سه گام ذیل است.

گام نخست: حذف داده‌های ناقص، بی‌کیفیت و مغشوش است. در این گام برخی از رکوردهای موجودی حذف می‌شود که فاقد محتوی هستند؛ یعنی باید برای هریک از مشتری‌های جامعه آماری مقدار هر سه ویژگی مدل (R,F,M) مشخص باشد، در غیر این صورت داده‌های این مشتری از جامعه آماری خارج می‌شود. لازم به ذکر است شاخص L,D برای مشتری نمونه انتخاب شده می‌تواند مقدار صفر بگیرد؛ به همین دلیل است که این دو شاخص در مدل به صورت جدا از سه معیار دیگر محاسبه می‌شود.

گام دوم: استخراج داده و ایجاد انباره داده‌ها است. هدف از این گام ایجاد یک انباره داده یکپارچه از تمامی مشتریان است. در این پژوهش منظور از انباره داده مخزنی از داده‌های جمع‌آوری شده در خصوص مشتریان بر مبنای شاخص‌های ارائه شده است.

گام سوم: این گام نرمال‌سازی داده‌ها است. باتوجه به اینکه داده‌های جمع‌آوری شده برای شاخص‌های مدل از یک جنس، مقیاس و جهت نیستند، باید داده‌های مذکور نرمال شوند؛ برای مثال R از جنس زمان و مقیاس روز است و جهت آن کاهشی است، یعنی R کمتر مطلوب‌تر است؛ در حالی که M از جنس مالی، مقیاس آن ریال و جهت آن افزایشی است، یعنی M بیشتر مطلوب‌تر است. برای نرمال‌سازی داده‌ها از روش Max-Min مطابق روابط ۵ الی ۹ برای هر مشتری استفاده شده است.

$$Ri^N = \frac{R_{Max} - Ri}{R_{Max} - R_{Min}} \quad \text{رابطه ۵}$$

$$Fi^N = \frac{Fi - F_{Min}}{F_{Max} - F_{Min}} \quad \text{رابطه ۶}$$

$$Mi^N = \frac{Mi - M_{Min}}{M_{Max} - M_{Min}} \quad \text{رابطه ۷}$$

$$Li^N = \frac{Li - L_{Min}}{L_{Max} - L_{Min}} \quad \text{رابطه ۸}$$

$$Di^N = \frac{D_{Max} - Di}{D_{Max} - D_{Min}} \quad \text{رابطه ۹}$$

در روابط بالا R_{Max} ، F_{Max} ، M_{Max} ، L_{Max} و D_{Max} نشان‌دهنده بیشترین مقادیر شاخص‌ها و R_{Min} ، F_{Min} و M_{Min} ، L_{Min} و D_{Min} بیانگر کمترین مقادیر شاخص‌ها هستند و R ، F ، M ، L و D نیز مقادیر اصلی شاخص‌ها را نشان می‌دهند. در نهایت Ri^N ، Fi^N ، Mi^N ، Li^N و Di^N بیانگر مقادیر نرمال‌شده شاخص‌ها برای مشتری نام است.

خوشه بندی براساس RFM-LD

باتوجه به اینکه در بخش اول پژوهش از داده‌کاوی برای فرایند کشف مدل خوشه‌بندی و ارزش‌های مجموعه داده‌های مدنظر استفاده می‌شود، برای پیاده‌سازی آن یک الگوی مشخص لازم است. یکی از الگوهای قوی در ادبیات موضوع، متدولوژی CRISP-DM^{۲۹} است. این متدولوژی از گام‌های شناخت داده، آماده‌سازی داده، مدل‌سازی، ارزیابی و توسعه تشکیل شده است. دنبال کردن این الگو به داده‌کاوی کمک می‌کند که با نگاه عمیق‌تری از گزارشات معمول داده‌ها به صورت منطقی تحلیل شود (چمپن و همکاران^{۳۰}، ۲۰۰۰)؛ بنابراین برای تفکیک مشتریان در گروه‌های همگن براساس ارزش شاخص‌های مدل، از روش خوشه‌بندی کا- میانگین (K-means) استفاده شده است. دستور کلی خوشه‌بندی K-میانگین در نرم‌افزار R به صورت رابطه ۱۰ به کار گرفته شده است.

رابطه ۱۰ `kmeans(x, centers, iter.max = 10, nstart = 1, algorithm = c("Hartigan-Wong", "Lloyd", "Forgy", "MacQueen"), trace=FALSE)`

المان‌های به‌کاررفته در این روش برای هر بار خوشه‌بندی به صورت زیر است.

X: ماتریس عددی از داده‌ها یا همان ماتریس شاخص‌های استانداردشده برای مشتریان (نمونه‌ها) است.

Centers: تعداد خوشه‌ها است که باید به صورت پارامتری به دستور اضافه شود.

iter.max: حداکثر تعداد تکرار مجاز برای انجام خوشه‌بندی است (شرط توقف).

nstart: چنانچه centers یک عدد باشد، تعداد مجموعه‌های تصادفی را نشان می‌دهد که باید انتخاب شود.

Algorithm: خوشه‌بندی به این روش برای اجرا از الگوریتم‌های متفاوتی استفاده می‌کند. این المان برای تابع `kmean` الگوریتم خاصی را مشخص می‌کند. این تابع به طور پیش فرض الگوریتم هارتیگان و ونگ را در خوشه بندی پیش می‌گیرد؛ اما ممکن است روش‌های دیگری (الگوریتم مک کوپین) استفاده شود که اغلب به سایر الگوریتم‌ها ترجیح داده می‌شود؛ اما به طور کلی الگوریتم هارتیگان و ونگ نسبت به بقیه بهتر عمل می‌کند.

Trace: تنها در روش هارتیگان و ونگ کاربرد دارد و گویای اعداد صحیح یا منطقی در پیشبرد الگوریتم را در نظر می‌گیرد (هارتینگ و وانگ^{۳۱}، ۱۹۷۵). برای تعیین تعداد خوشه در روش کا-میانگین از روش وارد^{۳۲} بهره گرفته شده است. در این روش همانند دیگر روش‌های خوشه‌بندی سلسله مراتبی، نخست هر یک از اشیاء

به صورت خوشه جداگانه‌ای در نظر گرفته می‌شوند. سپس در هر گام، همه ترکیب‌های دوتایی ممکن از ادغام دیده می‌شود و شاخصی به نام مجموع توان دوم خطا^{۳۳} محاسبه می‌شود. هر ترکیبی که خطای کمتری داشته باشد، برگزیده می‌شود و برپایه آن ادغام انجام می‌شود. روشن است که با هر بار مقدار حداقل خطا افزایش می‌یابد. این کار تا زمانی تکرار می‌شود که همه اشیاء با هم ادغام شوند و خوشه یکتایی حاصل شود (وارد^{۳۴}، ۱۹۶۳). گاهی از این روش به نام روش "کمترین واریانس" یاد می‌شود.

مدل‌سازی پژوهش

برای ساخت مدل ریاضی در بیان مسئله پژوهش نیازمند ارائه تعاریف و مفاهیم استفاده‌شده در محصولات، مفروضات و فرایندها ساخت مدل است که در ادامه به صورت گام به گام مطرح شده است.

گام ۱) تعریف محصولات و بیان مفروضات مدل: براساس قانون بانکداری بدون ربا که در سال ۱۳۶۲ به تصویب رسیده است، وظیفه اصلی سیستم بانکی کشور تجهیز و تخصیص منابع پولی است. بنابراین تمامی بانک‌ها و مؤسسات اعتباری مجاز هستند خدمات و محصولات خود را در راستای این دو وظیفه اصلی تعریف کنند. در وظیفه نخست بانک‌ها که امین آحاد جامعه و مشتریان خود هستند در جمع‌آوری و نگهداری پول (تجهیز منابع خود) از طریق ارائه بخشی از محصولات خود با عنوان طرح‌های سپرده‌ای سپرده‌های مشتریان را جذب می‌کنند (بهمند و بهمنی، ۲۰۰۶). وظیفه دوم عملیات بانکی، اعطای تسهیلات بانکی به مشتریان است. در مدل این پژوهش بانک محصولات خود را در دو گروه تسهیلاتی و سپرده‌ای به مشتریان عرضه می‌کند. گروه تسهیلاتی از ۹ محصول و گروه سپرده‌ای از ۴ محصول تشکیل شده است. ویژگی‌های هر کدام از محصولات ارائه‌شده به مشتریان به وسیله بانک در بازه زمانی اشاره‌شده در جدول ۱ آمده است.

جدول ۱- ویژگی‌های محصولات استفاده‌شده در مدل

نوع محصول	نماد	نرخ مؤثر درصد	ریسک درصد	مطلوبیت درصد	هزینه
تسهیلات ^(۱)	y1	۲۷	۳۸/۸	۰/۰۱	۰/۱۲
	y2	۲۳/۶	۱۱	۰/۳	۰/۰۳
	y3	۱۷/۵	۴/۹	۰/۱۶	۰/۰۳
	y4	۱۴/۳	۱/۷	۰	۰
	y5	۱۸/۳	۴/۵	۰/۰۹	۰/۰۲
	y6	۱۹/۸	۷/۳	۰/۰۷	۰/۰۲
	y7	۰	۲/۹	۰	۰/۰۲
	y8	۰	۲/۷	۰/۳۶	۰/۰۲
	y9	۲۰	۰	۰	۰/۰۱
سپرده‌ها	x1	۰	۰	۰/۳۱	۰
	x2	۰	۰	۰/۰۶	۰
	x3	۲۰	۰/۲	۰/۶۱	۱/۲
	x4	۲۲	۰/۰۰۲	۰/۰۲	۱/۲۲

در جدول ۱ محصولات در گروه تسهیلات و سپرده‌ها تفکیک شده‌اند. مطابق تعاریف به‌کاررفته در مجموعه قوانین و مقررات بانک مرکزی، منظور از نرخ مؤثر نرخ واقعی ارائه محصولات به مشتری به‌وسیله بانک، منظور از ریسک در گروه تسهیلات نسبت کل تسهیلات غیرجاری به جاری و در گروه سپرده‌ها نسبت مجموع دارایی‌های نقد به مجموع سپرده‌های مشتریان نزد بانک است. میزان مطلوبیت هر محصول، از مجموع نسبت تعدادی و مبلغی هر محصول در گروه خود حاصل شده و در نهایت هزینه هر واحد محصول در گروه تسهیلات ناشی از هزینه ذخیره‌گیری هر نوع وام و هزینه هر واحد از سپرده‌ها ناشی از هزینه سود پرداختی به آنها است.

گام ۲) ساخت مدل ریاضی: در ساخت مدل ریاضی پژوهش از دو نوع متغیر تصمیم، سه تابع هدف و پنج محدودیت استفاده شده است. متغیرهای تصمیم این مدل $x_{i,j}$ ، یعنی تخصیص سپرده λ م به خوشه λ م و $y_{i,j}$ ، یعنی تخصیص تسهیلات λ م به خوشه λ م است. در این مدل سه تابع هدف از دیدگاه بانک تعریف شده است. تابع هدف نخست رضایت یا مطلوبیت مشتریان از تخصیص محصولات به آنها بررسی شده است. تابع هدف دوم هزینه بانک ناشی از هر واحد تخصیص خدمات و محصولات به هر کدام از خوشه‌های مشتریان و تابع هدف سوم میزان ریسک ناشی از تخصیص هر واحد محصولات به مشتریان است. در تابع هدف نخست، مشتریان به‌دنبال کسب بیشترین مطلوبیت یعنی دریافت سود سپرده‌ای زیاد و تسهیلاتی با نرخ‌های کم هستند؛ درحالی‌که دومین تابع هدف به‌دنبال کاهش هزینه ناشی از پرداخت هزینه سود سپرده و کاهش هزینه ذخیره‌گیری تسهیلات با نرخ‌های زیاد است. در نهایت سومین تابع هدف به‌دنبال کاهش ریسک ناشی از تخصیص محصولات به مشتری برای بانک است. باتوجه به آنکه ضرایب R, F, M, L و D مبین یک تابع توزیع احتمالی مشخص هستند، هر کدام از توابع به صورت امید ریاضی تعریف می‌شوند.

محدودیت نخست براساس قوانین و مقررات بانک مرکزی هر بانک نمی‌تواند بیش از ۸۰ درصد منابع خود را به تسهیلات اختصاص دهد؛ بنابراین در هر خوشه نسبت مجموع تسهیلات اختصاص داده شده به مجموع سپرده‌های مشتریان نباید بیشتر از ۸۰ درصد باشد. به‌دلیل بهره‌مندی حداقل یک نوع سپرده به مشتریان یک خوشه، در محدودیت دوم مجموع سهم درصد هر یک از انواع سپرده‌های تخصیص داده شده به هر خوشه بیش از ۱۰ درصد است. هم‌چنین به‌دلیل مشابه در محدودیت سوم مجموع سهم درصد هر یک از انواع تسهیلات تخصیص داده شده به هر خوشه بیش از ۱ درصد است؛ یعنی مشتریان هر خوشه باید حداقل از یک نوع تسهیلات استفاده کنند. در محدودیت چهارم و پنجم مقدار متغیرهای تصمیم به‌صورت باینری (صفر و یک) است.

$$\begin{aligned} \text{Max. } Z_1 &= \sum_i \sum_j \text{Customer satisfaction} = \sum_i n_i \cdot \left(\frac{\bar{F}_i}{\bar{R}_i^w} \right) \sum_j (Q_{x_{i,j}} \cdot x_{i,j} + Q_{y_{i,j}} \cdot y_{i,j}) \\ \text{Min. } Z_2 &= \sum_i \sum_j \text{Cost of Bank} = \sum_i n_i \left((\bar{M}_i^w \sum_j C_{x_{i,j}} \cdot x_{i,j}) + (\bar{L}_i^w + \bar{D}_i^w) \sum_j C_{y_{i,j}} \cdot y_{i,j} \right) \\ \text{Min. } Z_3 &= \sum_i \sum_j \text{Risk of Bank} = \sum_i n_i \left((\bar{M}_i^w \sum_j R_{x_{i,j}} \cdot x_{i,j}) + (\bar{L}_i^w \sum_j R_{y_{i,j}} \cdot y_{i,j}) \right) \\ \text{s.t} \\ \frac{n_i \bar{L}_i^w \sum_j y_{i,j} \cdot y_{i,j}}{n_i \bar{M}_i^w \sum_j x_{i,j} \cdot x_{i,j}} &\leq 0.8 \rightarrow \frac{\bar{L}_i^w \sum_j y_{i,j} \cdot y_{i,j}}{\bar{M}_i^w \sum_j x_{i,j} \cdot x_{i,j}} \leq 0.8 \\ \forall i &= 1, \dots, n \quad \forall j = 1, \dots, m \end{aligned}$$

$$n_i \bar{M}_i^w \sum_j I_{x_{i,j}} \cdot x_{i,j} \geq 0.1 n_i \bar{M}_i^w \rightarrow \sum_j I_{x_{i,j}} \cdot x_{i,j} \geq 0.1 \quad \forall i = 1, \dots, n$$

$$n_i \bar{L}_i^w \sum_j I_{y_{i,j}} \cdot y_{i,j} \geq 0.01 n_i \bar{L}_i^w \rightarrow \sum_j I_{y_{i,j}} \cdot y_{i,j} \geq 0.01 \quad \forall i = 1, \dots, n$$

$$x_{i,j} = \{0, 1\}$$

$$y_{i,j} = \{0, 1\} \quad \forall i = 1, \dots, n \quad \forall j = 1, \dots, m$$

n_i : تعداد مشتریان خوشه نام.

\bar{F}_i : متوسط تراکشن‌های مشتریان خوشه نام.

\bar{R}_i^w : متوسط وزنی آخرین زمان مراجعه مشتریان خوشه نام.

\bar{M}_i^w : متوسط وزنی مانده سپرده مشتریان خوشه نام.

\bar{L}_i^w : متوسط وزنی تسهیلات مشتریان خوشه نام.

$Q_{x_{i,j}}$: مطلوبیت هر واحد تخصیص از محصول زام (اندیس x یعنی سپرده زام) به مشتریان خوشه نام.

$Q_{y_{i,j}}$: مطلوبیت هر واحد تخصیص از محصول زام (اندیس y یعنی تسهیلات زام) به مشتریان خوشه نام.

$C_{x_{i,j}}$: هزینه هر واحد تخصیص از محصول زام (اندیس x یعنی سپرده زام) به مشتریان خوشه نام.

$C_{y_{i,j}}$: هزینه هر واحد تخصیص از محصول زام (اندیس y یعنی تسهیلات زام) به مشتریان خوشه نام.

$R_{x_{i,j}}$: ریسک هر واحد تخصیص از محصول زام (اندیس x یعنی سپرده زام) به مشتریان خوشه نام.

$R_{y_{i,j}}$: ریسک هر واحد تخصیص از محصول زام (اندیس y یعنی تسهیلات زام) به مشتریان خوشه نام.

$I_{x_{i,j}}$: سهم درصد محصول زام (اندیس x یعنی سپرده زام) از کل سپرده‌ها برای مشتریان خوشه نام.

$I_{y_{i,j}}$: سهم درصد محصول زام (اندیس y یعنی تسهیلات زام) از کل تسهیلات برای مشتریان خوشه نام.

$x_{i,j}$: تخصیص سپرده زام به خوشه نام (اگر تخصیص یابد یک در غیر این صورت صفر).

$y_{i,j}$: تخصیص تسهیلات زام به خوشه نام (اگر تخصیص یابد یک در غیر این صورت صفر).

گام ۳) روش تبدیل اهداف چندگانه به یک تابع مطلوبیت: در بسیاری از مسائل چندهدفه، جواب بهینه وجود ندارد؛ زیرا در اکثر مواقع اهداف در تضاد با هم هستند و بهینگی یک هدف باعث دورشدن از اهداف دیگر می‌شود؛ بنابراین جواب بهینه در مدل‌های چندهدفه لزوماً مترادف با بهینه‌شدن تمامی هدف‌ها نیست. برای حل مدل‌های چندهدفه روش‌های مختلفی وجود دارد که جواب هر روش با روش دیگر لزوماً یکسان نیست؛ زیرا مفروضات هر روش و میزان مشارکت تصمیم‌گیرنده در فرایندهای حل متفاوت است. روش‌هایی از قبیل تبدیل تابع هدف به محدودیت، وزن‌دهی به اهداف، اولویت مطلق، روش معیار جامع، روش L-P متریک، برنامه‌ریزی آرمانی، برنامه‌ریزی سازشی و غیره در ادبیات موضوع دیده می‌شود (مومنی، ۲۰۱۲). در این پژوهش از روش برنامه‌ریزی سازشی برای تک‌هدفه‌کردن مدل استفاده شده است. در این روش، مدل چند هدفه با در نظر گرفتن هریک از توابع هدف به صورت مجزا حل و سپس مدل به صورت تک‌هدفه بازنویسی می‌شود به طوری که تابع هدف جدید به دنبال حداقل کردن اختلاف نرمالیزه هر تابع هدف با مقدار بهینه آن است. در مدل ارائه‌شده این پژوهش سه تابع هدف z_1

Z_2 و Z_3 وجود دارد که قبلاً معرفی شده‌اند. مدل با در نظر گرفتن هریک از این توابع به صورت مجزا حل شده است و مقادیر بهینه Z^*1 و Z^*2 به ترتیب برای تابع هدف نخست و دوم به دست می‌آید. سپس تابع هدف جدید به صورت زیر بازنویسی می‌شود (w_1 و w_2 بیانگر مقدار ضریب برنامه‌ریزی سازشی است به طوری که مجموع آنها برابر یک است). در این پژوهش اوزان تابع هدف یکسان لحاظ شده است. برای توابع هدفی که از نوع بیشینه‌سازی است، فرم تبدیل MOLP $_{35}$ به SOLP $_{36}$ به صورت رابطه ۱۱ است.

$$Z = \sum_{k=1}^2 W_k \times \frac{1}{H_k} \times Z_k \quad \text{رابطه ۱۱}$$

K : اندیس تابع هدف k ام است.

W_k : وزن تابع هدف k ام است که با توجه به اهمیت هر کدام از اهداف تعریف می‌شود.

H_k : حداکثر مقدار مطلوب طیف مقادیر تابع هدف k ام است.

در نهایت تابع هدف مدل به کاررفته پژوهش به صورت زیر است.

$$\begin{aligned} \text{Max. } Z = E & \left(\left(\frac{1}{H_Q} \right) \sum_i n_i \cdot \left(\frac{\bar{F}_i}{\bar{R}_i^w} \right) \sum_j (Q_{x_{i,j}} \cdot x_{i,j} + Q_{y_{i,j}} \cdot y_{i,j}) \right. \\ & - \left(\frac{1}{H_C} \right) \sum_i n_i \left((\bar{M}_i^w \sum_j C_{x_{i,j}} \cdot x_{i,j}) + (\bar{L}_i^w + \bar{D}_i^w) \sum_j C_{y_{i,j}} \cdot y_{i,j} \right) \\ & \left. - \left(\frac{1}{H_R} \right) \sum_i n_i \left((\bar{M}_i^w \sum_j R_{x_{i,j}} \cdot x_{i,j}) + (\bar{L}_i^w \sum_j R_{y_{i,j}} \cdot y_{i,j}) \right), \theta \right) \end{aligned}$$

H_Q : حداکثر میزان مطلوبیت تابع هدف نخست (مطلوبیت).

H_C : حداقل میزان هزینه تابع هدف دوم (هزینه).

H_R : حداقل میزان ریسک تابع هدف سوم (ریسک)

تجزیه و تحلیل داده‌ها و حل مدل.

θ : پارامتر توابع توزیع احتمال است.

در این پژوهش پس از آنکه مشتریان براساس پنج شاخص خوشه‌بندی شدند، توابع توزیع هر کدام از شاخص‌های مدل براساس ویژگی‌های تمامی مشتریان موجود در هر خوشه محاسبه شده است. با توجه به آنکه یکی از مهم‌ترین رویکردهای شبیه‌سازی استفاده از طراحی روش‌های آزمایش و انجام آزمایشات روی مدل مورد مطالعه است. در این رویکرد تمامی شرایط مختلف برای حل مسئله تحت عنوان سناریو در نظر گرفته شده، بررسی می‌شود. هم‌چنین در این پژوهش سناریوها (جواب‌ها) به کمک طراحی عاملی 2^k تولید می‌شود. از آنجاکه تخصیص یا عدم تخصیص یک محصول به یک خوشه متغیر تصمیم در نظر گرفته شده است از K به عنوان تعداد محصولات استفاده می‌شود ($k=13$)؛ اما به دلیل آنکه در عمل تولید $2^{(13 \times 7)}$ سناریو برای ۷ خوشه ممکن نیست از $k-p$ استفاده می‌شود ($p=10$)؛ بنابراین برای این پژوهش به کمک نرم افزار MINITAB ۸ سناریو تولید شده است. هر کدام از سناریوها در نرم افزار Arena ترسیم شده‌اند و پس از اجرا مقادیر متغیرها و تابع هدف محاسبه شده است تا ورودی‌های الگوریتم تبرید^{۳۷} شبیه‌سازی شده برای بهبود جواب حاصل شود.

بحث

داده‌های استخراج شده از انبار داده‌ای مشتریان یکی از بانک‌های خصوصی کشور طی بازه زمانی یک‌ساله ۱۳۹۵ استخراج و شاخص زمان محاسبه و استخراج داده‌های انتخاب شده در انتهای سال تنظیم شد (این شاخص در انبار داده با عنوان "تاریخ مؤثر" برای استخراج داده استفاده می‌شود). کل داده‌های درخور استفاده پس از عملیات پاک‌سازی برای تمامی مشتریان حقیقی بانک، ۳۱,۹۵۳ کد مشتری (حدود ۱۰ درصد) بود و تمامی متغیرهای مربوط به ۵ شاخص R, F, M, L و D در جدول ۲ استخراج شد.

جدول ۲- اطلاعات آماری داده‌های مرتب با شاخص‌های RFMLD

مشتریان حقیقی	کمترین مقدار	بیشترین مقدار	میانگین	انحراف معیار	میان
آخرین زمان مراجعه/ تازگی (R) (روز)	۰	۳۶۲	۲۶۱	۶۹	۲۸۵
تعداد تراکنش (F) (عدد)	۴	۲۴۰,۲۲۴	۲۵۴	۱,۶۴۲	۳۵
متوسط کمترین مانده سپرده (M) (ریال)	۵	۱۵۳,۰۴۳,۰۲۴	۴,۷۵۵,۵۵۹	۱۳,۷۹۹,۸۸۸	۱۸۹,۲۲۷
متوسط مانده جاری تسهیلات (L) (ریال)	۰	۲۰۰,۰۰۰,۰۰۰,۰۰۰	۲۳,۳۹۲,۳۲۸	۱,۴۳۹,۱۹۸,۵۳۴	۰
متوسط مطالبات (D) (ریال)	۰	۶,۸۲۵,۹۵۰,۶۸۵	۱,۳۰۷,۴۵۸	۴۷,۶۲۹,۴۴۷	۰

پس از بررسی مقدماتی نرمال بودن متغیرهای پژوهش و اثبات این ادعا که متغیرهای پژوهش نرمال نیستند و همچنین به دلیل حجم نمونه زیاد (بیشتر از ۲۰۰۰ نمونه) از آزمون کولموگروف-اسمیرنوف^{۳۸} استفاده شد. هنگام بررسی نرمال بودن داده‌ها فرض صفر مبتنی بر اینکه توزیع داده‌ها نرمال است در سطح خطای ۵ درصد تست شد. نتایج مربوطه در جدول ۳ آمده است. H_0 فرض توزیع داده‌های مربوط به هریک از متغیرهای نرمال است. در مقابل فرض H_1 ، توزیع داده‌های مربوط به هریک از متغیرهای غیرنرمال است.

جدول ۳- نتایج آزمون آزمون کولموگروف-اسمیرنوف

متغیرها	آماره آزمون	Df درجه آزادی	Sig. سطح معنی داری
R	۰/۱۸۷	31953	.000
F	۰/۴۳۹	31953	.000
M	۰/۳۶۵	31953	.000
L	۰/۴۹۴	31953	.000
D	۰/۴۸۹	31953	.000

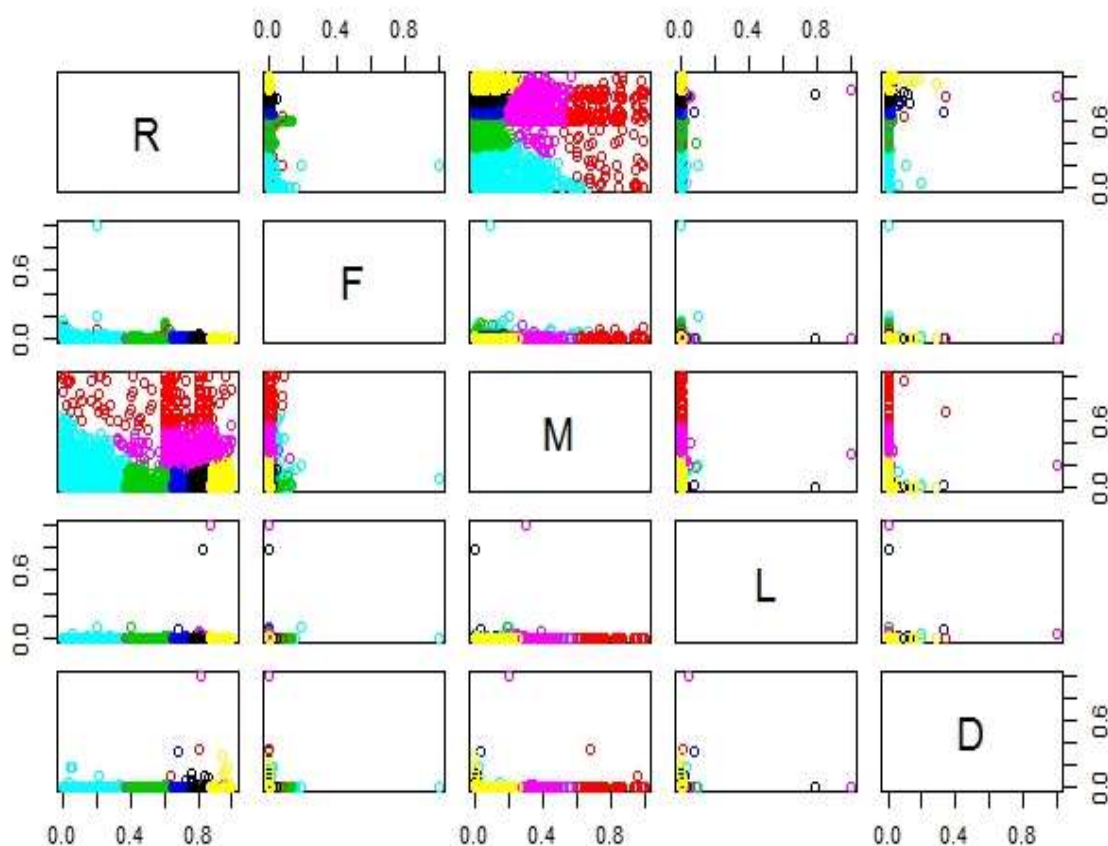
بنابراین با توجه به جدول ۳ و مقدار سطح معنی داری Sig که در آزمون انجام شده کمتر از ۵ درصد است، فرض نرمال بودن داده‌ها (H_0) برای تمامی متغیرهای پژوهش رد می‌شود و توزیع داده متغیرها نرمال نیست.

وزن کلیه شاخص‌ها به روش آنتروپی محاسبه شد و اوزان به دست آمده برای هر کدام از شاخص‌های R, F, M, L و D به ترتیب ۰/۰۰۳، ۰/۱۱۲، ۰/۱۱۵، ۰/۳۹۵ و ۰/۳۷۶ به دست آمد که نشان دهنده بالابودن میزان اهمیت متوسط وزنی تسهیلات دریافتی و متوسط وزنی معوقات به وسیله مشتریان است. خلاصه وضعیت خوشه‌بندی و تعداد اعضای هر خوشه در جدول ۴ نشان داده شده است.

باتوجه به منطبق روش وارد و با استفاده از شاخص سیلوئت (نیمرخ) تعداد خوشه موردانتظار ۷ خوشه به دست آمد که در شکل ۱ نمایش وضعیت هر خوشه (۷ خوشه رنگی) در مختصات هر یک از شاخص‌های R, M, F, L به تصویر کشیده شده است. لازم به ذکر است زمان اجرای الگوریتم برای حل با استفاده از نرم‌افزار آر استودیو نسخه ۳ روی یک سیستم کامپیوتری ۲ هسته‌ای ۲ گیگاهرتز دارای ۴ گیگابایت حافظه داخلی ۲ دقیقه و ۴۱ ثانیه محاسبه شده است.

جدول ۴- تعداد مشتریان در هر بار خوشه‌بندی با تعداد خوشه k

شماره خوشه	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
1	12,362	11,981	12,339	4,757	208	4,705	1,961	1,024	9,226
2	19,591	1,149	2,014	11,673	2,003	1,000	6,001	1,955	2,355
3	-	18,823	1,033	1,019	11,590	3,376	2,777	6,866	4,465
4	-	-	16,567	2,006	4,744	1,962	5,019	1,994	4,707
5	-	-	-	12,498	12,373	5,919	6,253	4,793	1,232
6	-	-	-	-	1,035	6,431	7,040	207	204
7	-	-	-	-	-	8,560	1,902	5,883	1,940
8	-	-	-	-	-	-	1,000	6,234	759
9	-	-	-	-	-	-	-	2,997	4,266
10	-	-	-	-	-	-	-	-	2,799
جمع	31,953	31,953	31,953	31,953	31,953	31,953	31,953	31,953	31,953
$\frac{\text{between_SS}}{\text{total_SS}}$	۴۳/۲	74	۸۳/۵	۸۵/۳	۸۸/۸	۹۰/۱	۹۰/۶	۹۱/۴	۹۱/۷

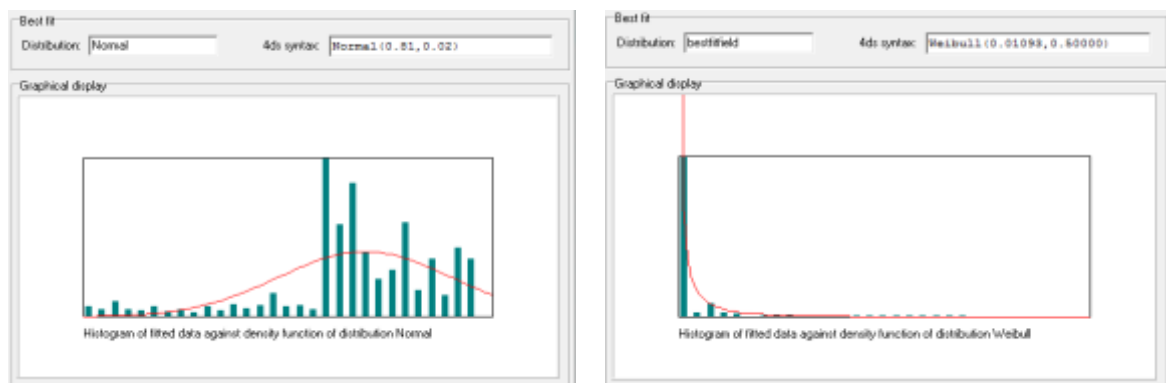


شکل ۱- نمایش وضعیت هر خوشه (۷ خوشه رنگی) در مختصات هر یک از شاخص‌های R, M, F, L, D

باتوجه به آنکه هرکدام از متغیرهای R, L, M, F, D متناسب با نوع فعالیت مشتری در زمان‌های مختلف تغییر می‌یابد، از دستور Auto Fit از طریق اتم ExcelActiveX_Read در نرم‌افزار ED استفاده و تابع توزیع هرکدام از شاخص‌ها برای هر خوشه محاسبه شد و در جدول ۵ نشان داده شده است.

جدول ۵- توابع توزیع شاخص‌های هر یک از خوشه‌ها

	R	F	M	L	D
خوشه نخست	Uniform (0/830,0/894)	Lognormal (0/0001212,0/0003037)	Lognormal (0/012,0/033)	Negexp (0/0000611)	Logistic (0/0001852,0/0024956)
خوشه دوم	Normal (0/69,0/14)	Lognormal (0/002,0/007)	Lognormal (0/42,0/2)	Negexp (0/0012)	Negexp (0/0005)
خوشه سوم	Uniform (0/89,1/00)	Logistic (0/00006,0/00014)	Lognormal (0/01,0/02)	Normal (0/000033,0/000177)	Negexp (0/0004)
خوشه چهارم	Beta (0/13,0/60,1/20)	Lognormal (0/01,0/03)	Weibull (0/06,0/50)	Negexp (0/0002)	Negexp (0/000336)
خوشه پنجم	Lognormal (0/69,0/03)	Weibull (0/0007,0/8)	Weibull (0/01609,0/50000)	Negexp (0/0000414)	Negexp (0/0000985)
خوشه ششم	Logistic (0/61,0/04)	Lognormal (0/00141,0/00384)	Weibull (0/027728,0/7)	Negexp (0/000056)	Logistic (0/0000456,0/0006003)
خوشه هفتم	Normal (0/81,0/02)	Lognormal (0/00039,0/00098)	Weibull (0/01093,0/50000)	Negexp (0/0001305)	Negexp (0/000215)



شکل ۲- توابع توزیع شاخص‌های R و M برای خوشه هفتم

در شکل ۲ برای نمونه توابع توزیع شاخص‌های R و M برای خوشه هفتم نمایش داده شده است. باتوجه به این شکل تابع توزیع شاخص R برای مشتریان خوشه هفتم دارای تابع توزیع نزدیک به نرمال با میانگین $0/81$ و واریانس $0/02$ است. هم‌چنین برای شاخص M دارای توزیع وایبل با پارامترهای $0/01093$ و $0/5$ است. پس از تعیین توابع توزیع متغیرها، مدل ریاضی به کمک شبیه‌سازی برای 100 هزار مشتری در 8 سناریو اجرا شد. جواب نزدیک به بهینه در بخش نتیجه‌گیری تشریح می‌شود.

نتیجه‌گیری

برای این پژوهش ابتدا براساس مدل تخصیص، تمام 8 سناریو مدل شبیه‌سازی یارانه‌ای به کمک نرم‌افزار Arena طراحی شده برای هر بار اجرا یک جواب موجه برای مسئله تولید می‌شود. درحقیقت نقش مدل شبیه‌سازی در این پژوهش تولید جواب‌های موجه به صورت تصادفی برای مسئله است؛ زیرا همه محدودیت‌های مسئله شبیه‌سازی می‌توانند در برنامه‌سازی کدنویسی شوند. با این کار براساس تابع هدف مدل ریاضی پژوهش، تعداد مشخصی

از جواب‌های موجه که مناسب‌ترین مقادیر تابع هدف را به خود اختصاص می‌دهند وارد الگوریتم تبرید می‌شوند و از آن به بعد برای الگوریتم تبرید با تولید و بهبود جواب‌های جدید فرایند بهینه‌سازی به اتمام می‌رسد. فرض صحیح در اینکار تولید جواب موجه ابتدایی (جمعیت اولیه) خوب و باکیفیت به جای جمعیت اولیه تصادفی است که می‌تواند زمان حل الگوریتم را تا حد زیادی کاهش دهد. ضمن آنکه با تولید جواب‌های شبیه‌سازی شده احتمال رسیدن به جواب نزدیک به بهینه در میان آنها افزایش می‌یابد. شرط توقف الگوریتم رشد بهبود جواب‌های همسایه کمتر از ۰/۰۰۱ در نظر گرفته شده است. در نهایت نتایج حل مدل ریاضی در جدول ۷ نمایش داده شده است.

جدول ۷- جواب نهایی مسئله

نوع توابع هدف								Min H _R	Min H _C	Max H _Q
توابع هدف								ریسک (Z3)	هزینه (Z2)	مطلوبیت (Z1)
مقدار بهینه Z برای هر کدام از توابع هدف								۴۷۷	۲۴۰۴	۲۷۷
مقدار تابع هدف شبیه‌سازی شده								۰/۶۶۱۱۳۴۱۷۴		
Xij7	Xij6	Xij5	Xij4	Xij3	Xij2	Xij1	Z3	جواب مسئله بهینه به کمک شبیه سازی		
1	1	1	1	0	0	1	y1			
1	1	0	1	0	0	0	y2			
1	1	1	1	1	1	1	y3			
1	1	1	1	1	1	1	y4			
1	1	1	1	1	1	1	y5			
1	1	1	1	0	1	0	y6			
1	1	1	1	1	1	1	y7			
1	1	1	1	1	1	1	y8			
1	1	1	1	1	1	1	y9			
1	1	1	1	1	1	1	x1			
1	1	0	1	1	1	1	x2			
1	1	0	1	0	0	1	x3			
1	1	1	1	0	1	1	x4			

همان‌طور که در حل نهایی مسئله نشان داده شده است، عدد یک برای X_{ij} نشان‌دهنده تخصیص محصول تسهیلاتی (y) و یا سپرده‌ای (x) i ام به هر کدام از ۷ خوشه زام بوده است. در مقابل عدد صفر نشان‌دهنده تخصیص ندادن محصول تسهیلاتی (y) و یا سپرده‌ای (x) i ام به هر کدام از ۷ خوشه زام است. با توجه به جواب به دست آمده و در صورت داشتن شرایط قانونی برای هر مشتری نتیجه می‌شود، مشتریان خوشه‌های ۶ و ۷ می‌توانند از کلیه وام‌ها و سپرده‌های بانک برخوردار شوند. از طرف دیگر برای مشتریان خوشه سوم افتتاح سپرده‌های جاری و قرض الحسنه توصیه می‌شود و سپرده‌های کوتاه یا بلندمدت توصیه نمی‌شود. هم‌چنین برای این خوشه از مشتریان پرداخت وام‌های مشارکتی یا مضاربه‌ای توصیه نمی‌شود.

مهم‌ترین نوآوری این پژوهش طراحی و به‌کارگیری مدل ریاضی چندهدفه و استفاده از روش خوشه‌بندی در ارائه خدمت و محصولات بانک به مشتریان است که در حل آن از بهینه‌سازی از طریق شبیه‌سازی استفاده شده است. از محدودیت‌های این پژوهش ناتوانایی ابزارهای شبیه‌سازی در ترسیم و حل تمامی حالات محتمل (سناریوهای بیشتر) و حل مدل برای آنها است. پیشنهاد آتی این پژوهش، توسعه مدل ریاضی از بُعد مشتری است که پس از حل

آن بانک قادر به تصمیم‌گیری برای ارائه خدمت و محصول به تک‌تک مشتریان خود است و هم‌زمان توابع هدف در مناسب‌ترین حالات خود قرار می‌گیرند. درنهایت با استفاده از یک مدل بتوان پارامترهای مربوط به هرکدام از محصولات را برای مشتری جدید مراجعه‌کننده به بانک تنظیم کرد. نوآوری و وجه تمایز اصلی این پژوهش با سایر مطالعات به شرح ذیل است.

۱- اکثر مطالعات انجام‌شده به صورت پراکنده تمرکزشان روی روش، مدل حل مسئله، کوتاه‌کردن زمان و ابعاد حل مسئله با به‌کارگیری الگوریتم‌های مختلف است؛ درحالی‌که این پژوهش به صورت کاربردی و علمی در جهت حل یک مسئله در دنیای واقعی شکل گرفته و تمام تمرکز آن روی ایجاد مدل جامعی است که تا حد زیادی عوامل واقعی موجود در فضای مسئله را پوشش دهد و بتواند آن را حل کند، به طوری که با اضافه کردن دو شاخص تسهیلات دریافتی مشتریان و معوقات ایجاد شده برای بانک به مدل معروف RFM و توسعه آن باعث محاسبه دقیق‌تر ارزش مشتریان بانک و استفاده از آن در خوشه‌بندی آنها شده است (سجادی و همکاران، ۲۰۱۵).

۲- در مطالعات پیشین معمولاً بدون ذکر دلیل با استفاده از روش‌های مختلف مشتریان خوشه‌بندی شده‌اند؛ درحالی‌که در این پژوهش روش‌های خوشه‌بندی و جایگاه آن در داده‌کاوی و دلایل استفاده از روش به‌کاررفته تشریح شده است.

در اکثر مطالعات انجام‌شده پس از انجام خوشه‌بندی مشتریان، مسئله رها شده و تنها با انجام تحلیل‌های موردی و کوتاه به اتمام رسیده است؛ درحالی‌که در این پژوهش مسئله با انجام فرایند خوشه‌بندی شروع و نتایج آن به‌عنوان ورودی به مدل برای به‌کارگیری آن استفاده شده است و توانسته نیاز اصلی پژوهش را به صورت جامع پوشش دهد. ۳- این پژوهش با در نظر گرفتن فرایندی جامع ابتدا مشتریان بانک را براساس ویژگی‌های اساسی و مهم، خوشه‌بندی کرده است؛ سپس با شناسایی محصولات بانکی قابل عرضه به مشتری، مدلی طراحی شده است که اهداف اساسی بانک در عرصه خدمات و محصولات را به مشتری پوشش می‌دهد. درنهایت با استفاده از مفاهیم موجود در روش‌های بهینه‌سازی از طریق شبیه‌سازی مسئله حل شده است؛ درحالی‌که این فرایند جامع در مطالعات پیشین به صورت متمرکز مطرح نشده است.

از محدودیت‌های این پژوهش ناتوانایی ابزارهای شبیه‌سازی در ترسیم و حل مدل با تمام اجزاء و مفروضات بیشتر موجود در دنیای واقعی است. علاوه بر آن محدودیت در استفاده از نسخه‌های آموزشی نرم‌افزارهای به‌کاررفته در پژوهش و محدودیت کار با استفاده از رایانه به‌کاررفته وجود دارد.

پیشنهاد آتی این پژوهش، توسعه مدل ریاضی از بعد مشتری اشاره است. درواقع پس از حل آن بانک قادر به تصمیم‌گیری برای ارائه خدمت و محصول به تک‌تک مشتریان خود است و هم‌زمان توابع هدف در مناسب‌ترین حالات خود قرار گیرند. درنهایت با استفاده از یک مدل بتوان پارامترهای مربوط به هرکدام از محصولات را برای مشتری جدیدی تنظیم کرد که برای نخستین بار به بانک مراجعه می‌کند. پیشنهاد دیگر برای سایر پژوهشگران آنکه با استفاده از سایر الگوریتم‌های داده‌کاوی برای خوشه‌بندی و الگوریتم‌های فراابتکاری دامنه جواب‌های مدل ریاضی کاوش و نتایج آن مقایسه شود. علاوه بر این پیشنهاد می‌شود با استفاده از تکنیک‌های داده‌کاوی، تحلیل رفتار مشتریان در بازه‌های زمانی مختلف رصد و تغییر سلاقی آنها در استفاده از خدمات و محصولات بانک پایش شود.

References

- Abiodun, R. (2017). "Development of Mathematical Models for Predicting Customers Satisfaction in the Banking System with a Queuing Model Using Regression Method". *American Journal of Operations Management and Information Systems*, 2(2), 86-91.
- Afsar A, houshdar mahjoub R, minaei B.(2014) "Customer credit clustering for Present appropriate facilities". *Management Research in Iran*, 17 (4), 1-24 .
- Akbariasl, R., & Bashli, M. (2014) "Banking Services Marketing", *Ettihad Publishing, Tehran, Iran*. 81-92.
- Alborzi, M., Mohammad Pourzarandi, M. E., Khanbabayi, M. (2010). "Using Genetic Algorithm in optimizing decision trees for credit scoring of banks customers". *Journal of Information Technology Management*, 2(4), 23-38.
- Athanassopoulos, A. D. (2000). "Customer satisfaction cues to support market segmentation and explain switching behavior", *Journal of Business Research*, 47 (3), 191-207.
- Azizi, S., Balaghi Inalou, M. (2016). "Segmentation of Mobile Banking Users Based on Expectations: A Clustering Technique". *Journal of Production and Operations Management*, 7(2), 217-234. doi: 10.22108/jpom.2016.21103.
- Bahmand, M., & Bahmani, M. (2006). "Internal Banking (Supply of Money Resources)", *Iranian Institute of Banking Publisher, Tehran, Iran*, 48-50.
- Baradara, V., Farokhi, Z. (2014). "Customer Segmentation in Banking Industrial by RFCM Model development", *Journal of Brand Management*, 1(2), 135-154.
- Buttle, F., (2004) "Customer Relationship Management: Concepts and Tools", *Elsevier Butterworth Heinemann*.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). "CRISP-DM 1.0: Step-by-step data mining guide". *SPSS Incc*.
- Chen, Q., Zhang, M., & Zhao, X. (2017). "Analysing customer behaviour in mobile app usage". *Industrial Management & Data Systems*, 117, 425-438.
- Ghorbanpour A, tallai G, panahi M. (2015) "Clustering Customers of Refah Bank Branches Using Combination of Genetic Algorithm and C- Means in Fuzzy Environment". *Organizational Resources Management Researchs.*; 5 (3), 153-168.
- Hartigan, J. (1975), "Clustering algorithms", *Wiley, New York*.
- Izadi, B., Ranjbarian, B., Ketabi, S., Nasiri Mofakham, F. (2016). "A Comprehensive Approach for Market Segmentation and Customer Classification Using Data Mining methods and alternative Linear Programming". *Journal of Production and Operations Management*, 7(1), 1-22. doi: 10.22108/jpom.2016.20909
- Keiningham, T. and Aksoy, L. and Bejou, D., (2006) "Approaches to measurement and management of customer value", *Journal of relationship marketing*, 5(2), 37-54.
- Kotler, P., (1994) "Marketing management: Analysis, planning, implementation, and control", *New Jersey: Prentice-Hall*.
- Kumar, M., Vishnu Chaitanya, M. and Madhavan, M. (2012) "Segmenting the banking market strategy by clustering" , *International Journal of Computer Applications*, Vol. 45, No. 17, pp. 10-15.
- Mishar, A., (2009) "Customer relationship management: Implementation process perspective", *Acta polytechnic hungarica.*, 6(4), 83-99.

- Momeni, M. (2012). "Data Clustering (Cluster Analysis)", *Danesh Negar Publisher, Tehran, Iran*, 37-38
- Nazemi, Jamshid., Jafari, Pejman., Hashemi, Hamed.(2012). "Mining the Retail Banking Customers Characteristics Using Data Mining Techniques". *Journal of Marketing Management* .2(7), 12, 21-35.
- Ngai, E.W.T. ,(2005) "Customer relationship management research (1992-2002): An academic literature review and classification", *Marketing intelligence Planning*, Vol. 23,pp.582-605.
- Safari Kahreh, M. and Safari Kahreh, Z.,(2012) "An empirical analysis to design enhanced customer lifetime value based on customer loyalty: evidences from Iranian banking sector", *Iranian journal of management studies*, 5(2), 145-167.
- Shahrabi, Jamal., Zolghadr, Ali.,(2011) "Advanced data mining :concepts algorithms", *Soroush Gita Publisher*. 130:142.
- Singh, S., & Singh, S. (2016). "Accounting for risk in the traditional RFM approach". *Management Research Review*, 39(2), 215-234.
- Sohrabi, Babak., Khanlari, Amir., Ajourlo, Nasibeh.(2011)."A Model to Determine Customer Lifetime Value (CLV) in Banking Industry". *Management Research in Iran*. 15 (1) :224-239
- Taghavifard, M., Khajvand, S. & Najafi, E. (2011). "Customer clustering Saderat Bank of Iran by using data mining". *Improvement Management Studies*, 67(21), 197-200
- Sajjadi, K., Khatami-Firuzabadi, M. A., Amiri, M., Sadaghiani, J. S. (2015).A developing model for clustering and ranking bank customers. *InternationalJournal of Electronic Customer Relationship Management*, 9(1), 73-86.
- Taghavifard, M., Khajvand, S. (2013). "Customer Segmentationof Bank Customers (Case: SaderatBank of Iran)". *Journal of Business Administration Researches*, 5(9), 39-64.
- Thomas, J. S. (2001). "A methodology for linking customer acquisition to customer retention", *Journal of Marketing Research*, 38 (2), 262-268.
- Wang, H. F. & Hong, W. K. (2006). "Managing customer profitability in a competitive market by continuous data mining", *Industrial Marketing Management*, 35 (6), 715-723.
- ward j. h. jr. (1963) "hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, 58 (301), 236-244.
- Wu, Hsin-Hung; Chang, En-Chi and Lo, Chiao-Fang,(2009) "Applying RFM model and K-means method in customer value analysis of an outfitter", *International Conference on Concurrent Engineering*, New York.
- Zabkowski, T. (2016). "RFM approach for telecom insolvency modeling". *Kybernetes*, 45(5), 815-827.

¹- Abiodun

²- Varun Kumar

³- Key Account

⁴- Thomas

⁵- Kotler

⁶- Athanassopoulos

⁷- Chen

⁸- Ngai

⁹- Mishar

¹⁰- Wang

¹¹- Singh

¹²- Zabkowski

¹³- Customer Life time Value

¹⁴- Net Present Value

- 15- Share of Wallet
- 16- Markov Chain
- 17- Past Customer Value
- 18- Return on Investment
- 19- Recency, Frequency, Monetary
- 20- Battel
- 21- Keiningham
- 22- k-Means
- 23- Wu
- 24- Recency
- 25- Frequency
- 26- Monetary
- 27- Loan
- 28- Deferred
- 29- Cross Industry Standard Process for Data Mining
- 30- Champan
- 31- Hartigan, Wong
- 32- Ward
- 33- Sum of squares error index
- 34- Ward
- 35- Maly Objective Linear Programing
- 36- Single Objective Linear Programing
- 37- Simulated Annealing
- 38- Kolmogorov-Smirnov