Industrial Management Perspective

Print ISSN: 2251-9874, Online ISSN: 2645-4165
Volume 10, Serial Number 40, Winter 2021, Pages 41-63 (Original Article)
DOI: 10.52547/JIMP.10.4.41

Clustering and Prediction Model of Customer Lifetime Value (Case Studies: IRAN National Center for Numbering Goods and Services)

Fateme Nabizade[*], Saeed Rouhani[**]

Abstract

One of the important topics in this field is Customer Lifetime value (CLV) that means how much profit a customer generates in his lifetime for a company. The main goal of this research is presenting a model for clustering and predicting customer lifetime value and customer evaluation in IRAN National Center for Numbering Goods and Services. In this research, 74,385 records of members at a specified time interval were used (from 2011 to 2017). Members are classified by CRISP methodology, resulting in the presentation of a model for predicting them. At first, members are classified into 7 clusters by RFM and K-Means. Next, each cluster is rated by CLV. Next, hidden patterns are discovered inside the data and various segments of members are then predicted through classification algorithms. Finally, the algorithms are evaluated and the final report is prepared. The results of this study exhibit the member's behavior in each cluster on the organization's services and membership or extended subscription and also future customer behavior are unveiled. This research helps the managers to come up with marketing strategies to keep loyal members and attract or remove inactive members by analyzing the clusters.

Keyword: Customer Lifetime Value; Data Mining; RFM; Clustering; Predict.

Received: Jul. 27, 2020; Accepted: Sep. 12, 2020.
* Masters, Mehralborz University.
** Associate Professor, University of Tehran.
E-mail: srouhani@ut.ac.ir

چشم‌انداز مدیریت صنعتی

شاپای چاپی: ۹۸۷۴-۲۲۵۱، شاپای الکترونیکی: ۴۱۶۵-۲۶۴۵

سال دهم، شماره ۴۰، زمستان ۱۳۹۹، ص‌ص ۴۱ – ۶۳ (**نوع مقاله: پژوهشی**)

DOI: 10.52547/JIMP.10.4.41

مدل خوشه‌بندی و پیش‌بینی ارزش طول عمر مشتری (مورد مطالعه: مشتریان مرکز شماره‌گذاری کالا و خدمات ایران)

فاطمه نبی‌زاده*، سعید روحانی**

چکیده

یکی از مباحث مهم در زمینه حفظ مشتریان و چگونگی رفتار با آن‌ها، ارزش طول عمر مشتری (CLV) است . هدف از این پژوهش، طراحی مدلی برای خوشه‌بندی و پیش‌بینی طول عمر مشتریان و همچنین ارزیابی مشتریان در مرکز شماره‌گذاری کالا و خدمات ایران است. در این پژوهش اطلاعات ۷۴۳۸۵ عضو این سازمان در بازه زمانی ۱۳۹۰-۱۳۹۶ دریافت شد. مشتریان توسط تکنیک داده‌کاوی CRISP طبقه‌بندی شده و درنهایت مدلی برای پیش‌بینی آن‌ها طراحی شد. ابتدا اعضا توسط مدل RFM و الگوریتم K-Means به ۷ طبقه دسته‌بندی شده و سپس هر طبقه توسط روش محاسبه ارزش طول عمر مشتریان رتبه‌بندی شد. در ادامه توسط الگوریتم‌های رگرسیون لجستیک، درخت تصمیم و شبکه‌های عصبی، الگوهای پنهان بین داده‌ها و بخش‌های مختلف مشتریان کشف شدند. نتایج این پژوهش، رفتار مشتریان هر یک از خوشه‌ها را در خدمات مرکز و همچنین مدل رفتار مشتریان آتی را نشان داده است. این پژوهش با تحلیل خوشه‌ها به مدیران در ارائه راهبردهای بازاریابی، حفظ اعضای وفادار و جذب یا حذف اعضای غیرفعال، یاری می‌رساند. در پژوهش حاضر تعداد خوشه مناسب برای مشتریان ۷ عدد است؛ همچنین در پیش‌بینی کلاس مشتریان عملکرد شبکه‌های عصبی با دقت ۹۹/۵۶ درصد نسبت دیگر الگوریتم‌ها بهتر بوده است.

کلیدواژه‌ها: ارزش طول عمر مشتری؛ داده‌کاوی؛ RFM؛ خوشه‌بندی؛ پیش‌بینی.

تاریخ دریافت مقاله: ۱۳۹۸/۰۵/۰۵، تاریخ پذیرش مقاله: ۱۳۹۹/۰۶/۲۲.

* کارشناسی ارشد، دانشگاه مهر البرز.

** دانشیار، دانشگاه تهران (نویسنده مسئول).

E-mail: srouhani@ut.ac.ir

۱. مقدمه

در عصر جدید یکی بزرگ‌ترین چالش‌های سازمان‌های مشتری‌محور، شناخت مشتریان، بخش‌بندی آن‌ها به گروه‌های مختلف، رتبه‌بندی آن‌ها و پیش‌بینی مشتریان جدید است؛ به‌طوری‌که تضمین بلندمدتی برای تداوم کسب‌وکار به‌وجود آورد. درنتیجه سازمان باید شناخت درستی از مشتریان خود داشته باشد و نیازها و خواسته‌هایشان را پیش‌بینی کند و با استفاده از اطلاعات به‌دست‌آمده بهره‌وری خود را افزایش دهد. مدیریت ارتباط با مشتری[1] راه‌حلی است که به سازمان‌ها کمک می‌کند به شیوه‌ای سازمان‌یافته، ارتباط با مشتریان را مدیریت کنند. یکی از چالش‌های مهم در حوزه مدیریت ارتباط با مشتری در سازمان‌های مشتری‌محور، شناخت مشتریان و درک تفاوت و طبقه‌بندی آن‌ها است [۲۰]. سازمان‌ها و شرکت‌ها علاقه‌مند هستند مشتریان وفادار خود را شناسایی کنند و نسبت به آن‌ها رفتار و یا خدماتی مناسب ارائه دهند؛ همچنین مشتریان غیروفادار را با راهبردهایی به مشتریان وفادار تبدیل کنند و یا خدماتی مناسب با ارزش آن‌ها ارائه دهند و در غیر این‌صورت آن‌ها را از دایره خدمات حذف کنند. برای مدیریت درست ارتباط با مشتری می‌توان از روش محاسبه ارزش طول عمر مشتری[2] استفاده کرد. نتایج ارزش طول عمر مشتری محاسبه‌شده برای بخش‌های مختلف می‌تواند برای توضیح استراتژی‌های بازاریابی و فروش توسط شرکت به‌کار رود [۲۰]. برخی از مدل‌های ارزش طول عمر مشتری برای تقسیم‌بندی بازار و تخصیص منابع بازاریابی برای کسب، نگهداری و فروش متقابل مفید هستند [۲۸]. چالش اصلی اندازه‌گیری طول عمر مشتری، توسعه مدلی برای پیش‌بینی جریان درآمد آینده است که هر مشتری به شرکت ارائه می‌دهد [۲۶].

سؤال‌های مطرح شده در پژوهش حاضر به شرح زیر است:

ـ چگونه می‌توان اعضای «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» را خوشه‌بندی نمود؟

ـ چگونه ارزش طول عمر اعضای «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» را محاسبه کنیم؟

ـ چگونه می‌توان با استفاده از مفهوم ارزش طول عمر مشتری، ارزش اعضای مرکز را سنجید؟

ـ چگونه می‌توان مدل پیش‌بینی طول عمر مشتریان را در «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» ارائه کرد؟

هدف از انجام پژوهش حاضر، طراحی مدل پیش‌بینی طول عمر مشتریان در «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» است؛ به‌طوری‌که برای این منظور باید معیارهای CLV بر اساس پارامترهای مؤثر مشتریان، وزن و اهمیت هر یک از شاخص‌های مؤثر در دسته‌بندی مشتریان و تکنیک و ابزارهای مناسب CLV مشخص شود.

---

1. Customer Relationship Management (CRM)
2. Customer Life Time Value (CLV)

در بخش بعدی مفاهیم به‌کار رفته در این پژوهش معرفی و سپس پژوهش‌های پیشین در این زمینه مرور خواهد شد. در بخش بعد روش پژوهش ارائه می‌شود و سپس به تجزیه‌وتحلیل داده‌های به‌دست‌آمده پرداخته خواهد شد. درنهایت نیز نتایج حاصل از تحلیل داده‌ها نشان داده می‌شود.

بیان مسئله. یکی از چالش‌های مهم در حوزه مدیریت ارتباط با مشتری در سازمان‌های مشتری‌محور، شناخت مشتری و درک تفاوت و طبقه‌بندی آن‌ها است [۲۳]. سازمان‌ها و شرکت‌ها علاقه‌مند هستند مشتریان وفادار خود را شناسایی کنند و نسبت به آن‌ها رفتار و یا خدماتی مناسب ارائه دهند و همچنین مشتریان غیروفادار را با راهبردهایی تبدیل به مشتریان وفادار کنند و یا خدماتی مناسب با ارزش آن‌ها ارائه دهند و در غیر این صورت آن‌ها را از دایره خدمات حذف کنند.

در این پژوهش کاربردی، با هدف خوشه‌بندی مشتریان با استفاده از روش‌های داده‌کاوی و تخمین ارزش طول عمر مشتری، «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» به‌عنوان مورد مطالعه انتخاب شد. روش پیشنهادی پژوهش برگرفته از روش کریسپ[1] است. «این مرکز، خدماتی چون صدور کد ملّی (شناسایی کالاهای ملّی) و خدمات کد بین‌المللی[2] را ارائه می‌دهد. یکی از مسائلی که این سازمان با آن روبه‌رو است، شناسایی مشتریان و طول عمر آن‌ها است تا توسط آن بتواند مشتریان وفادار و غیروفادار خود را شناسایی و به‌صورت استراتژیک نسبت به آن‌ها تصمیم‌گیری کند. در این پژوهش تلاش شده است تا با استفاده از تکنیک خوشه‌بندی و الگوهای درون‌داده طبق داده‌های استخراج‌شده، اعضای سازمان موردمطالعه بر اساس محاسبه ارزش طول عمر مشتری طبقه‌بندی شده و مدلی برای پیش‌بینی طول عمر مشتری جدید ارائه شود. اهمیت موضوع طبقه‌بندی زمانی مشخص می‌شود که بر اساس آمار رسمی مرکز، در بازه زمانی سال‌های ۱۳۹۰–۱۳۹۶ تعداد ۷۴۳۸۵ نفر از اعضای این سازمان ۴۸۰۹۹۳ کد GS1 و ۱۸۱۳۷۴۲ ایرانکد دریافت کرده‌اند. چگونگی حل این مسئله به‌گونه‌ای که سازمان بتواند برنامه‌های استراتژیک آینده را بر اساس شناخت مناسبی که از بخش‌های مختلف مشتریان به‌دست‌آمده آورده است، تبیین کند؛ به‌طوری‌که سود شرکت افزایش یابد و درباره ارزش طول عمر مشتریان جدید پیش‌بینی‌های مناسبی صورت پذیرد.

1. Cross-Industry Standard Process for Data Mining (CRISP-M)
2. The Global Language of Business (GS1)

۲. مبانی نظری و پیشینه پژوهش

تعاریف مفهومی

بازاریابی. بازاریابی در راستای دستیابی به اهداف موردنظر همواره برای انجام مبادلات مطلوب در بازارهای هدف در تلاش بوده است. دانش بازاریابی با تغییرات بنیادی در مفاهیم فلسفی خود سعی کرده است تا نیازها و خواسته‌های مشتریان را به نحو مطلوب‌تر و مؤثرتر پاسخ دهد [۳۰]. عوامل بازاریابی از طریق شبکه‌های تلفن و یا اینترنت می‌توانند به‌صورت دائم در سطحی گسترده در بازار حضور یابند و نیازها و خواسته‌های مشتری را از زبان خودش بشنوند و به آن‌ها واکنش سریع نشان دهند؛ همچنین می‌توانند مشتریان را به نحوه دلخواه گروه‌بندی کنند و متناسب با روحیات و نیازهای هر گروه، اطلاع‌رسانی، پرسشگری و حتی گفت‌وگو کنند و برای مراسم خاص هدایا ارسال نمایند. بازاریابی مکانیزه فاصله بین شرکت و مشتری را از بین می‌برد و رابطه شرکت با مشتری را شفاف و سازنده می‌کند.

در بازاریابی ابتدا باید هویت مشتریان مشخص و خصوصیات هر یک از آن‌ها به‌طور مجزا بررسی شود. هرچه میزان ارتباط برقرارشده بین سازمان‌ها و مشتریانشان نزدیک‌تر و صمیمی‌تر باشد، به همان نسبت می‌توانند آن‌ها را به سازمان خود وابسته کنند. برقراری این ارتباط صمیمی نیازمند کسب اطلاعات دقیق از مشتریان است.

بازاریابی و تبلیغات بخش مهم اقتصاد است و تأثیر بسیاری بر ادراک مشتری از درک زندگی، فرهنگ و جهان دارد. برخی بر این باورند که بازاریابی باعث رشد اقتصادی و افزایش سطح رفاه عمومی می‌شود [۲]. از انواع بازاریابی می‌توان به «بازاریابی مستقیم» اشاره کرد. بازاریابی مستقیم نوعی از بازاریابی است که مستقیم با مشتری ارتباط برقرار می‌شود. در این روش افرادی که بازاریابی می‌شوند به‌طور مستقیم از بازار هدف انتخاب می‌شوند. نقطه مقابل بازاریابی مستقیم، «بازاریابی غیرمستقیم» نام دارد. در این نوع از بازاریابی به‌صورت غیرمستقیم سعی در جذب مشتری می‌شود؛ مانند تبلیغات بیلبوردی که در فضای عمومی صورت می‌گیرد. بعد از دسته‌بندی مشتریان، میزان تأثیر بازاریابی مستقیم بر هر یک از طبقات متناسب با خدمت موردنظر ارزیابی می‌شود. امروزه نیز بازاریابی به‌عنوان نخستین گام از استراتژی‌های مشتری‌محور موردتوجه قرار می‌گیرد [۱۰].

جایگاه مشتری در بازاریابی. یکی از واژه‌های پرکاربرد در سطح بازارهای داخلی و خارجی لغت «مشتری» است. تعریف‌های مختلفی برای این واژه وجود دارد. برای مثال، مشتری فرد یا افرادی هستند که کالا و خدمتی را دریافت می‌کنند یا مشتری بزرگ‌ترین سرمایه هر مؤسسه است [۱۹]. در تعریف‌های دیگر، مشتریان به‌عنوان مهم‌ترین منبع درآمدزایی هر سازمان به دو نوع تقسیم شده‌اند: ۱. مشتریان درونی که شامل همکاران فعلی سازمان هستند که در تولید کالا

یا ارائه خدمات نقش دارند؛ ۲. مشتریان بیرونی شامل افراد حقیقی و حقوقی خارج از سازمان هستند که محصول یا خدمت ارائه‌شده توسط سازمان را دریافت می‌کنند. در این پژوهش مفهوم لغت مشتری درواقع به «مشتری بیرونی» اطلاق می‌شود [۳]. نقطه آغاز بازاریابی، مشتری است. همان افرادی که سبب پویایی و کسب‌وکار می‌گردند، فرآیند بازاریابی دقیقاً از جایی شروع می‌شود که مشتری وارد چرخه‌ی کسب‌وکار می‌شود و نه زمانی که محصول آماده فروش است. درحقیقت مشتری، ارزش ایجادشده توسط کالا یا خدمات را خریداری می‌کند.

مدیریت روابط مشتریان الکترونیکی. گذر از اقتصاد و ازمیان‌رفتن مرزهای جغرافیایی برای کسب‌وکار و به‌تبع آن شدت‌یافتن رقابت، موجب شده است تا مشتری به‌عنوان رکن اساسی و محور اصلی فعالیت‌های سازمان مطرح شود. امروزه تنها رضایت‌مندی مشتریان برای ماندگاری آن‌ها در سازمان کافی نیت و باید در عین رضایت‌مندی از وفاداری آنان نیز اطمینان حاصل شود [۵]. مدیریت ارتباط با مشتری، مجموعه‌ای از روش‌هایی است که یک دیدگاه محکم، منسجم و یکپارچه از مشتریان در گستره کل کسب‌وکار فراهم می‌کند تا اطمینان حاصل شود هر مشتری بالاترین سطح خدمت را دریافت می‌کند [۲۵]. تقسیم‌بندی مشتریان به لحاظ رفتار به‌وجودآمده در هر بنگاه اقتصادی عامل اصلی در مدیریت ارتباط با مشتریان و همچنین بازاریابی هدفمند به‌شمار می‌رود [۱۸]. در جریان اجرای مدیریت ارتباط مؤثر با مشتری، فناوری اطلاعات نقش محوری در توانمندسازی سازمان‌ها به‌منظور حداکثرکردن سودآوری از طریق تعیین هدف‌های دقیق‌تر برای بخش‌های بازار و گره‌های درون آن‌ها دارد. فناوری اطلاعات با جمع‌آوری داده‌های مرتبط با مشتریان و سپس ایجاد بستر اجرای استراتژی‌های مناسب ارتباط با مشتری می‌تواند به یاری مدیران سازمان بشتابد [۲۷].

نسخه جدید CRM یا همان eCRM، اصولاً برخاسته از مدیریت روابط مشتری است؛ ولی بر شخصی‌سازی، فناوری‌های بازاریابی مستقیم برای فروش و ارائه خدمات متمایز به بخش‌های کوچک بازار تأکید بیشتری دارد. eCRM این امکان را برای مشتریان فراهم می‌کند تا با برقراری ارتباط مؤثرتر با شرکت، ظرفیت‌ها خود را به‌عنوان مشتری، بیشتر آشکار کنند تا شرکت و مشتری هر دو از این رابطه منتفع شوند [۸]. eCRM بر روی بستر CRM و با توجه مضاعف به فناوری‌های تقسیم‌بندی، تجزیه‌وتحلیل اطلاعات مربوط به مشتری، ارتباطات چندکاناله، تعاملات یک‌به‌یک و ارائه خدمات به بخش‌های خاص‌تری از بازار، رشد می‌کند [۱۵].

بر اساس نظر رومانو[1] (۲۰۰۱)، eCRM پرداختن به جذب و نگهداری مشتریان ارزشمند از نظر اقتصادی از طریق کانال‌های جدید و حذف آن‌هایی است که منافع کمی را برای شرکت به

1. Romano

دنبال دارند. وی همچنین بیان می‌دارد که eCRM با سیستم اطلاعات مدیریت و دیگر علوم مانند علوم رایانه، بازاریابی و روان‌شناسی ارتباط نزدیکی دارد و پابه‌پای این علوم به رشد و پیشرفت ادامه خواهد داد [۲۶]. استفاده از فناوری اطلاعات در مدیریت ارتباط با مشتریان مزایای زیادی ازجمله افزایش وفاداری مشتری، بازاریابی مؤثرتر، بهبود خدمات به مشتری و پشتیبانی از وی و کارایی بالاتر و کاهش هزینه‌ها دارد [۱۵].

ارزش طول عمر مشتری. به‌طورکلی بیش از ۳۰ سال پیش کاتلر[1] برای نخستین بار مفهوم ارزش مشتری را بدین‌صورت تعریف شد: «ارزش حال جریان سود آینده موردانتظار در طول افق یک دوره زمانی مشخص طی زمان ارتباط با مشتری» [۱۱].

ارزش دوره عمر مشتری عبارت است از: ارزش فعلی جریان منافعی که شرکت در صورت ازدست‌ندادن مشتری خود می‌تواند از او به‌دست آورد؛ بنابراین تلاش در راه حفظ و نگهداری مشتریان از نظر اقتصادی مقرون‌به‌صرفه بوده و موردتوجه شرکت‌ها است. در عصر مشتری‌محوری، ارزش مشتری، سلاحی استراتژیک در جذب و نگهداری مشتریان است [۱۴]. محاسبه ارزش طول عمر مشتری به شرکت کمک می‌کند تا بداند که چقدر می‌تواند در حفظ مشتری، سرمایه‌گذاری کند تا بازده مثبت سرمایه‌گذاری به‌دست آید؛ به‌عبارت‌دیگر CLV به شرکت کمک می‌کند تا به‌جای رفتار یکسان با همه مشتریان، با هر مشتری بر اساس سهم وی به‌طور متفاوتی رفتار کند [۳۱].

مدل RFM. یکی از مدل‌های ساده و درعین‌حال قدرتمند در اجرای CRM و تخمین ارزش مشتری، مدل RFM است [۵]. این مدل از سه بُعد مربوط به داده‌های مبادلاتی مشتریان برای تحلیل رفتار آن‌ها استفاده می‌نماید. شاخص‌های این مدل به‌صورت زیر تعریف می‌شوند [۴]:

تازگی مبادله (R): این شاخص بر فاصله زمانی بین آخرین خرید صورت‌گرفته توسط مشتری تا پایان دوره خاص (پایان محدوده زمانی موردبررسی) اشاره دارد. کمتربودن این فاصله نشان‌دهنده بالابودن ارزش این شاخص در مدل است.

تعداد تکرار مبادله (F): این شاخص تعداد مبادلاتی را نشان می‌دهد که یک مشتری در یک دوره زمانی خاص انجام داده است. بیشتربودن تعداد مبادلات نشان‌دهنده بالابودن ارزش این شاخص در مدل است.

ارزش پولی مبادله (M): این شاخص نشان‌دهنده مقدار پولی است که یک مشتری در یک دوره زمانی خاص برای مبادلات، صرف کرده است. بیشتربودن مقدار پول صرف‌شده نمایانگر بالابودن

1. Kotler

ارزش این شاخص در مدل است. این مدل یک مدل رفتارمحور است و به بخش‌بندی مشتریان و بررسی و پیش‌بینی رفتار مشتریان در آینده می‌پردازد.

پیشینه پژوهش. خواجوند و تارخ (۲۰۱۱)، پژوهشی در حوزه بانکداری انجام دادند که چارچوبی برای تقسیم مشتریان در بخش‌های همگن، محاسبه ارزش طول عمر مشتری در بخش‌های مختلف بر اساس تجزیه‌وتحلیل RFM و درنهایت مدلی برای پیش‌بینی هر بخش از مشتریان از روش سری‌زمانی فراهم می‌کند [۱۶].

خواجوند و همکاران (۲۰۱۱)، دو روش برای بخش‌بندی مشتریان و محاسبه ارزش طول عمر مشتری برای هر بخش ارائه کرده‌اند: ۱. روش RFM و ۲. روش پیشنهادی RFM پیشرفته با یک پارامتر اضافی (تعداد اقلام). نتایج این روش‌ها نشان داد که افزودن تعداد شمارش به‌عنوان یک پارامتر جدید در روش RFM هیچ نتیجه‌ای برای خوشه‌بندی ندارد [۱۷]. لی و همکاران[1] (۲۰۱۱) از مدل LRFM برای دسته‌بندی مشتریان در صنعت نساجی تایوان استفاده کردند؛ سپس به تحلیل بخش‌ها پرداختند [۲۱]. هیزوروگلو و سنگول (۲۰۱۲)[2] از سه رویکرد RFM، SOW و PCV برای طبقه‌بندی استفاده کردند. این رویکردها بر اساس ویژگی‌های اساسی مشتریان و ارزیابی یکسری از مدل‌های پیش‌بینی ارزش طول عمر مشتری ارائه شده‌اند [۱۳].

چنگ[3] (۲۰۱۲)، چارچوبی برای تعیین ارزش آینده مشتری در یک شرکت تعمیر و نگهداری خودرو در تایوان ارائه داد که با دیدگاه مدل CLV سنتی متفاوت بوده و دارای سه روش برای پیش‌بینی است: گروه نخست شامل یک مدل رگرسیون لجستیک و یک مدل درخت تصمیم‌گیری برای تخمین احتمال ریزش مشتری و برای پیش‌بینی بیشتر طول عمر مشتری در آینده است. گروه دوم شامل تجزیه‌وتحلیل رگرسیون برای شناسایی متغیرهای قطعی و یک زنجیره مارکوف برای مدل‌سازی است. گروه سوم شامل دو شبکه عصبی برای پیش‌بینی سود حاصل از یک مشتری تحت رفتارهای خرید مختلف است [۵].

دورسون و کبر[4] (۲۰۱۶)، بر روی پروفایل مشتریان سه هتل پنج ستاره در آنتالیا (ترکیه) با تحلیل RFM تمرکز کردند و بر اساس این مدل به تجزیه‌وتحلیل هر یک از طبقه‌بندی‌ها پرداختند [۹]. کریستی و همکاران[5] (۲۰۱۸)، ابتدا تحلیل RFM را بر روی داده‌های تراکنشی مشتریان اجرا کردند؛ سپس مشتریان با استفاده از الگوریتم‌های K-means و فازی C-means خوشه‌بندی شدند. در این پژوهش همچنین یک ایده جدید برای انتخاب مراکز اولیه خوشه در

1. Li & etal
2. Hiziroglu & Sengul
3. Cheng
4. Dursun & Caber
5. Christy & etal

الگوریتم K-means معرفی شده است [۶]. آنیتها و پاتیل[1] (۲۰۱۹)، پژوهشی بر پایه RFM در صنعت خورده‌فروشی انجام دادند و در آن برای طبقه‌بندی مجموعه داده‌ها از الگوریتم K-Means استفاده کردند. در این پژوهش تعداد و ارزیابی خوشه‌ها با استفاده از شاخص سیلوهت برای الگوریتم خوشه‌بندی K-Means صورت گرفت [۱]. قداده و عبدالله[2] (۲۰۱۸)، با استفاده از مجموعه داده‌های TIC CRM (مدیریت ارتباط با مشتری شرکت بیمه)، الگوریتم‌های ویژه K-Means و SOM را بررسی کردند. آن‌ها از ترکیب دو الگوریتم K-means به همراه شاخص دیویس بولدین[3] و SOM به‌طور هم‌زمان استفاده کردند. [۲۵]. مونالیزا و همکاران[4] (۲۰۱۹)، مشتریان شرکت LWC را توسط مدل RFM بررسی کردند. آن‌ها از الگوریتم خوشه‌بندی فازی C-means و تحلیل پورتفولیوی مشتریان بهره گرفتند. آن‌ها از مدل PCI برای اعتبارسنجی استفاده کردند و رتبه‌بندی با AHP برای یافتن ارزش طول عمر مشتری صورت گرفت. نتایج این پژوهش نشان داد که مشتریان شرکت LWC بر اساس مقدار CLV به‌دست‌آمده به سه خوشه‌ی سوپراستار (رتبه اول)، معمولی (رتبه دوم) و خفته (رتبه سوم) طبقه‌بندی می‌شوند [۲۲]. تارخ و اسماعیلی (۲۰۱۹)، مدل جدیدی برای محاسبه CLV در پزشکی طراحی کردند. آن‌ها از روش K-means به‌منظور خوشه‌بندی و از روش درخت تصمیم برای تحلیل خوشه‌ها و درنهایت از روش مارکوف برای پیش‌بینی رفتار بیمار استفاده کردند. در این پژوهش بر اساس زنجیره مارکوف توسعه‌یافته، فرمولی برای مدل CLV به‌دست آمد [۲۹].

هلت و همکاران[5] (۲۰۱۹) ، به مقایسه مدل RFM و RFM/P برای هر محصول پرداختند. نتایج این پژوهش نشان داد که داده‌های محصول می‌تواند اطلاعات مفیدی را برای مدیریت دارایی‌های بازاریابی و برآورد دقیق CLV فراهم کنند؛ همچنین می‌تواند خطای پیش‌بینی ارزش پایه مشتری را کاهش دهد [۱۲]. کاوودار و فرهتوسمانوکلوب[6] (۲۰۱۸)، مدلی برای برآورد ارزش طول عمر مشتری (CLV) در صنعت هواپیمایی معرفی کردند. آن‌ها در ابتدا با استفاده از اطلاعات پرواز مشتریان و سپس با افزودن داده‌های شبکه اجتماعی، مدل رگرسیون چندگانه را تهیه و سپس دو مدل به‌دست‌آمده را با هم مقایسه کردند. نتایج نشان داد که با افزودن عوامل اجتماعی به مدل پایه، هم دقت مدل و هم توانایی پیش‌بینی آن می‌تواند به‌طور بالقوه بهبود یابد و می‌تواند با مدل‌های موجود در حوزه‌های مختلف تطبیق داده شود [۳]. دانشور و همکاران (۲۰۲۰) با ترکیب الگوریتم K-Means و تکنیک پرامتی روش جدید خوشه‌بندی چندمعیاره را برای رفع نقایص روش P2CLUST ارائه دادند. آن‌ها با هوشمندسازی عملگر جهش، الگوریتم

1. Anitha & Patil
2. Qadadeh & Abdallah
3. Davies Bouldin
4. Monalisa & etal
5. Heldt & etal
6. Çavdar & Ferhatosmanoğlu

ژنتیک را ارتقا داده و یک الگوریتم ژنتیک ابتکاری معرفی کرده‌اند. نتایج نشان داد که الگوریتم خوشه‌بندی P2CLUST با استفاده از الگوریتم فراابتکاری مبتنی بر جمعیت ژنتیک برای حل مسئله خوشه‌بندی یک روش کارا، اما نیازمند زمان زیادی است [۷]. خاتمی و همکاران (۲۰۱۸)، بر اساس مدل RFMLD و الگوریتم K-Means مشتریان بانک را خوشه‌بندی و سپس مدل چندهدفه تخصیص خدمات به هر خوشه را ارائه کردند. برای بهبود جواب‌ها، سناریوهای مختلف توسط الگوریتم SA طراحی شد و هر یک از بخش‌ها تجزیه‌وتحلیل شدند [۱۸]. موتمنی و همکاران، مدلی ترکیبی از الگوریتم‌های تحلیل مؤلفه‌های مستقل، یادگیری منیفلد[1]، خوشه‌بندی K-Means و رگرسیون بردار پشتیبان برای پیش‌بینی میزان فروش در صنعت کاشی و سرامیک ارائه کردند. نتایج نشان می‌دهد که مدل پیشنهادی با درنظرگرفتن پیش‌پردازش داده‌ها (کم‌کردن اغتشاشات داده‌ها) از دقت بالاتری برخوردار است [۲۴].

جدول ۱. خلاصه مرور مبانی نظری

پژوهشگر/ سال	RFM	K-Means	رگرسیون	شبکه عصبی	درخت تصمیم	سیستم کدینگ
خواجوند و تارخ [۱۶]	✓					
خواجوند و همکاران [۱۷]	✓					
لی و همکاران [۲۱]	✓					
هیزوروگلو و سنگول [۱۳]	✓					
چنگ [۵]			✓	✓	✓	
دورسون و کبر [۹]	✓					
کریستی و همکاران [۶]	✓	✓				
آنیتها و پاتیل [۱]	✓	✓				
قداده و عبدالله [۲۵]		✓				
مونالیزا و همکاران [۲۲]	✓					
تارخ و اسماعیلی [۲۹]		✓			✓	
هلت و همکاران [۱۲]	✓					
کاوداروفرهتوسمانوکلوب[۳]			✓			
دانشور و همکاران [۷]		✓				
خاتمی و همکاران [۱۸]		✓				
موتمنی و همکاران [۲۴]		✓				
پژوهش حاضر	✓	✓	✓	✓	✓	✓

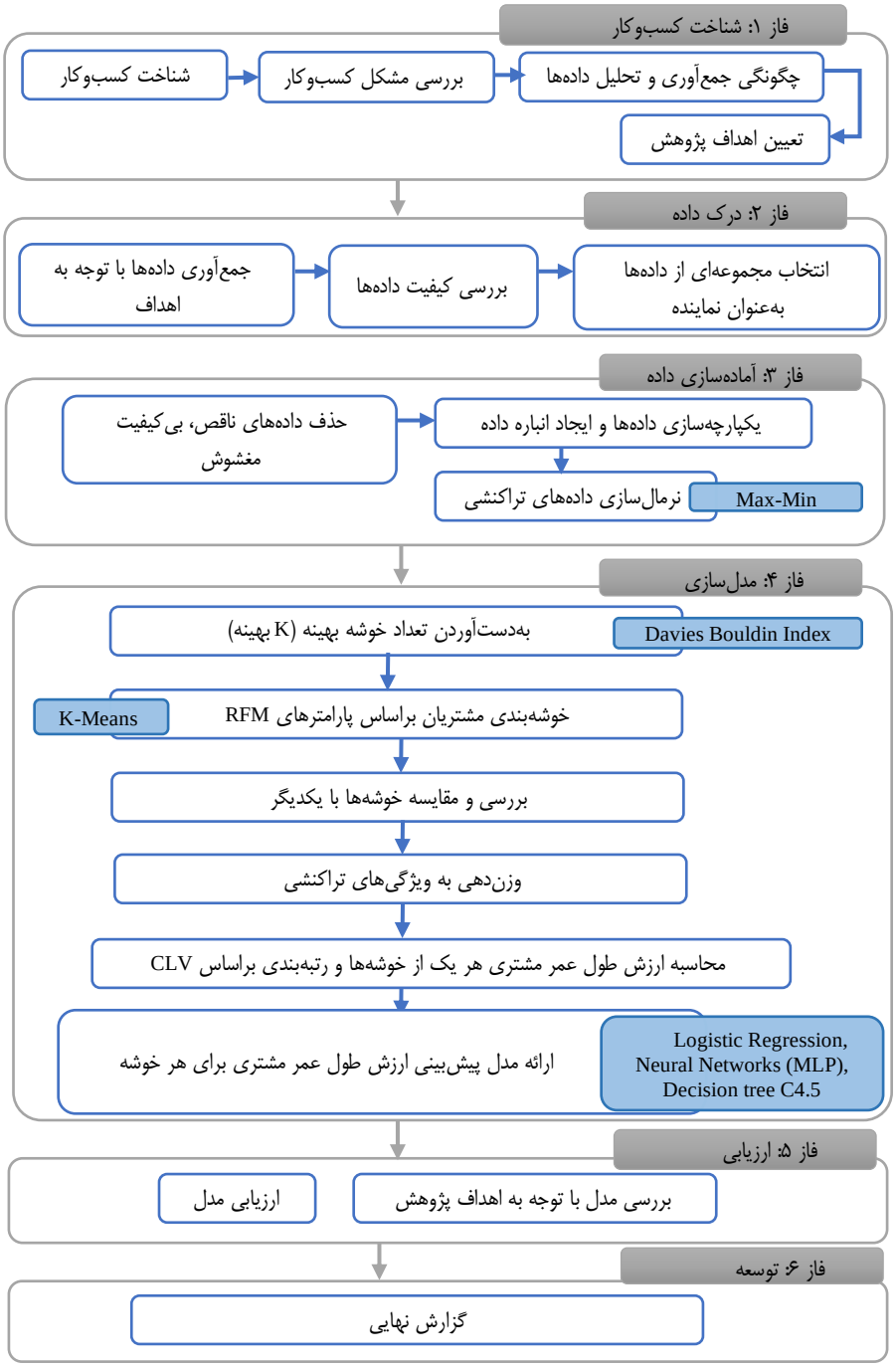
1. Manifold

با توجه به پیشینه پژوهش، عمده پژوهش‌های صورت‌گرفته پیرامون مفهوم ارزش طول عمر مشتری یا بر روش‌های یادگیری با نظارت متمرکز بودند و یا بر روش‌های یادگیری نظارت‌نشده. حال آنکه توسعه الگوهای ترکیبی مبتنی بر به‌کارگیری از هر دو رویکرد یادگیری با نظارت و بدون نظارت می‌تواند دارای مزایای زیادی در استفاده از الگوریتم های داده کاوی برای بررسی ارزش طول عمر مشتری باشد.

علی‌رغم پژوهش‌های زیاد صورت‌گرفته در استفاده از داده‌کاوی برای ارزش طول عمر مشتری بر روی پایگاه‌های داده‌های موجود در کسب و کارهای مختلف، مطالعه ساختاری و هدفمند در رابطه با سیستم‌های کدینگ کالا و خدمات صورت نگرفته است؛ به‌طوری‌که بیشتر پژوهش‌های انجام‌شده در این زمینه درباره مشتری‌های سیستم‌های بانکی و سایر کسب‌وکارهای مالی انجام شده است. با توجه به اهمیت سیستم‌های کدینگ و تنوع و توسعه روزافزون آن‌ها در سراسر جهان و اهمیت استفاده از این سیستم‌ها در ردیابی کالا و خدمات، ضرورت توسعه الگوهای مبتنی بر داده‌کاوی برای بررسی ارزش طول عمر مشتری در این حوزه بسیار زیاد است. نوآوری این پژوهش ارائه چارچوبی برای بخش‌بندی و محاسبه ارزش طول عمر مشتری در حوزه کدینگ کالا و خدمات است؛ همچنین پیش از این پژوهشی با درنظرگرفتن شاخص‌های مدنظر این پژوهش صورت نگرفته است. استفاده از الگوهایی برای پیش‌بینی CLV مشتریان آتی در این حوزه کسب‌وکار یکی دیگر از نوآوری‌های پژوهش حاضر است که بر اساس آن عملکرد سه روش دسته‌بندی‌کننده شامل شبکه‌های عصبی چندلایه، درخت تصمیم و رگرسیون با یکدیگر مقایسه شده و نتایج آن گزارش شده است.

۳. روش‌شناسی پژوهش

روش‌های مختلفی برای اجرای پروژه‌های داده‌کاوی ارائه شده‌اند؛ اما یکی از قدرتمندترین روش‌ها، روش CRISP است [۲۰]. این روش به‌عنوان یک روش فرآیندی، چرخه حیات یک پروژه داده‌کاوی را دربرمی‌گیرد. CRISP همچنین یک روش چرخشی است که شش مرحله عملیات داده‌کاوی را به‌طور کامل هدایت می‌کند. این شش مرحله شامل شناخت کسب‌وکار، شناخت داده، پیش‌پردازش، مدل‌سازی، ارزیابی مدل و توسعه مدل است. شکل ۱، مراحل عملیات داده‌کاوی مربوط به این پژوهش را نشان می‌دهد.



شکل ۱. مراحل پژوهش

مراحل پژوهش

شناخت کسب‌وکار. سازمان موردبررسی «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» است. این مرکز خدماتی همچون ارائه کُد ملّی و کدهای GS1 کالاها و خدمات را در اختیار اعضای خود قرار می‌دهد. موضوع این مطالعه کدهای ملّی و GS1 و عضویت/ تمدید اشتراک ارائه‌شده توسط این سازمان است. با توجه به موضوع این پژوهش، سازمان‌ها، شرکت‌ها و افراد با پرداخت حق عضویت می‌توانند یکی از اعضای ایرانکد محسوب شوند و در صورت نیاز می‌توانند سالانه اشتراک خود را تمدید کنند. مسئله رتبه‌بندی اعضا و شناسایی اعضای جدید، یکی از مسائلی که مرکز با آن مواجه بوده و در این پژوهش به آن پرداخته شده است. ایجاد سیستم بخش‌بندی و رتبه‌بندی مشتریان بر اساس عوامل مؤثر بر میزان ارزش مشتریان می‌تواند راهگشای حفظ مشتریان سودآور و فعال‌کردن اعضای غیرفعال این سازمان و تبیین برخی فعالیت‌های ارتباط با مشتری باشد.

درخواست و دریافت داده‌ها در یک مرحله صورت پذیرفت. اطلاعات درخواستی شامل اطلاعات اعضا و سرویس‌های دریافتی توسط هر یک از آن‌ها در زمان‌های مختلف است که طی هفت جدول با عناوین اعضا، اطلاعات اعضا، خدمات مرکز، جزئیات هر خدمت، پرداختی‌های عضو، خدمات اصلی ایرانکد اعضا و خدمات GS1 اعضا به‌دست‌آمده است. هدف این پژوهش، ارائه مدلی برای پیش‌بینی ارزش طول عمر مشتری است. ایجاد یک سیستم رتبه‌بندی، نیازمند استفاده از تکنیک‌های داده‌کاوی است. این پژوهش گروه متجانسی از مشتریان را موردتحلیل قرار داده است تا هم قابل‌فهم‌تر و هم به‌صرفه باشد.

درک داده. یکسری داده برای به‌دست‌آوردن پارامترهای مؤثر از جدول‌های ذکرشده در بالا واکشی شده است که در جدول ۲، مشاهده می‌شود.

جدول ۲. جداول جمع‌آوری‌شده

نام جدول	نام پارامترها
اعضا	کد عضو
ایرانکد	تاریخ ایجاد – ایرانکد
کد GS1	تاریخ ایجاد – کد GTIN
جزئیات اطلاعات اعضا	نام عضو – آخرین موجودی
تعرفه اعضاء	تاریخ تحویل – هزینه سالانه
خدمت	عنوان
تعرفه خدمت	تاریخ خدمت – مبلغ خدمت

سپس داده‌های به‌دست‌آمده از نظر درستی و پرت‌نبودن بررسی می‌شود. در مرحله بعد به‌دلیل بالابودن تعداد رکوردهای اطلاعات مشتریان و تراکنش‌های آنان، گزارش‌گیری بسیار زمان‌بر و مشکل بوده و نیاز به گرفتن گزارش‌ها در چند مرحله بود. پس از دریافت جدول‌ها توسط نرم‌افزار SQLSERVER، بانک اطلاعاتی یکپارچه‌ای شامل اعضا و تراکنش‌های آن‌ها طی یک دوره شش‌ساله (بازه زمانی سال‌های ۱۳۹۰– ۱۳۹۶) استخراج شد.

آماده‌سازی داده. در گام نخست برخی از پارامترهای موجود که دارای اطلاعات مفقوده زیاد بودند یا با سایر اطلاعات سازگاری نداشتند، حذف شدند.

در گام بعدی یک انباره داده یکپارچه بر اساس عضو از هفت جدولی که در فاز یک مطرح شد، به‌دست آمد. درنهایت اطلاعات به‌دست‌آمده مربوط به ۷۴۳۸۵ عضو است که در این مرحله شامل اطلاعات به‌دست‌آمده اطلاعات اصلی اعضا، فاصله آخرین تراکنش تا آخرین تاریخ تراکنش اعضا (روز) (R، تازگی تراکنش)، تعداد تراکنش‌های هر عضو (F، تعداد دفعات خرید) و جمع هزینه‌های مشتری در طول شش سال برای هر خدمت (M، حجم پولی خرید) است (RFM) (ابتدای سال ۱۳۹۰ تا انتهای سال ۱۳۹۶).

یادآوری می‌شود که تمام اطلاعات استخراج‌شده، به تفکیک هر یک از خدمات سازمان است.

جدول ۳. جزئیات داده‌های استخراج‌شده

نام پارامتر	توضیحات
RNC	تعداد روزهای گذشته از آخرین تراکنش دریافت کد ایرانکد تا زمان آخرین خرید این کد در کل اعضا
FNC	تعداد ایرانکد دریافتی در بازه زمانی یادشده
MNC	کل مبلغ پرداختی برای دریافت ایرانکد در بازه زمانی یادشده
RGS1	تعداد روزهای گذشته از آخرین تراکنش دریافت کد GS1 تا زمان آخرین خرید این کد در کل اعضا
FGS1	تعداد کد GS1 دریافتی در بازه زمانی یادشده
MGS1	کل مبلغ پرداختی برای دریافت کد بین‌المللی در بازه زمانی یادشده
RMT	تعداد روزهای گذشته از آخرین تراکنش تمدید اشتراک و یا عضویت تا زمان آخرین تمدید/ عضویت در کل اعضا
FMT	تعداد دفعات تمدید اشتراک و یا عضویت در بازه زمانی یادشده
MMT	کل مبلغ پرداختی برای تمدید اشتراک و یا عضویت در بازه زمانی یادشده

در این پژوهش داده‌های به‌دست‌آمده به روش Max-Min که در رابطه‌های ۱ تا ۳، قابل‌مشاهده است، نرمال می‌شود؛ سپس به کمک روش دیوس بولدین[1] تعداد خوشه‌های بهینه تعیین و در مرحله بعد اعضا بر اساس پارامترهای تحلیل RFM تطبیق‌یافته خوشه‌بندی می‌شوند. در گام بعدی به بررسی و مقایسه خوشه‌ها پرداخته شده و سپس هر یک از پارامترهای RFM وزن‌دهی می‌شود. در ادامه ارزش طول عمر مشتری برای هر یک از خوشه‌های موجود محاسبه و خوشه‌ها بر اساس ارزش CLV آن‌ها رتبه‌بندی می‌شوند. درنهایت با استفاده از الگوهای پیشین و مدل ارزش طول عمر مشتری در هر خوشه و انجام سعی و خطا، الگوهای سفارشی به‌دست می‌آید.

یادآوری می‌شود که تمام این عملیات در نرم‌افزار متلب ۲۰۱۸ انجام شده است.

رابطه (۱) $$R' = \frac{R_{max} - R}{R_{max} - R_{min}}$$

رابطه (۲) $$F' = \frac{F - F_{min}}{F_{max} - F_{min}}$$

رابطه (۳) $$M' = \frac{M - M_{min}}{M_{max} - M_{min}}$$

مدل‌سازی. این فاز شامل عملیات مدل‌سازی بوده که در ارتباط با داده‌های دریافت‌شده است. این فاز شامل هفت گام است:

۱. تعیین تعداد خوشه‌های بهینه به کمک روش دیوس بولدین؛

۲. خوشه‌بندی اعضا بر اساس پارامترهای تحلیل RFM تطبیق‌یافته؛

۳. بررسی و مقایسه خوشه‌ها با یکدیگر؛

۴. محاسبه وزن پارامترهای مؤثر در خوشه‌بندی؛

۵. محاسبه ارزش طول عمر مشتری برای هر یک از خوشه‌های موجود و رتبه‌بندی خوشه‌ها بر اساس ارزش CLV آن‌ها؛

۶. به‌دست‌آوردن الگوهای سفارشی با استفاده از الگوهای پیشین و مدل ارزش طول عمر مشتری در هر خوشه و انجام سعی و خطا.

۷. ارائه مدلی برای پیش‌بینی طول عمر مشتری بر اساس معیارهای سفارشی به روش طبقه‌بندی (درخت تصمیم، رگرسیون لجستیک و شبکه‌های عصبی چندلایه).

ارزیابی. برای ارزیابی مدل‌ها از شاخص‌های دقت، بازخوانی و صحت استفاده شده است.

1. Davies Bouldin Index

ـ ارزیابی مدل یا مدل‌های به‌دست‌آمده از فاز مدل‌سازی از لحاظ کیفیت و اثربخشی پیش از استفاده از نتایج آن‌ها؛
ـ تعیین اینکه آیا مدل به اهداف ترسیم‌شده در فاز نخست دست‌یافته است یا خیر؛
ـ احراز اینکه ابعاد کسب‌وکار به‌قدر کفایت موردتوجه قرارگرفته‌اند؛
ـ اتخاذ تصمیم در خصوص استفاده از نتایج داده‌کاوی.

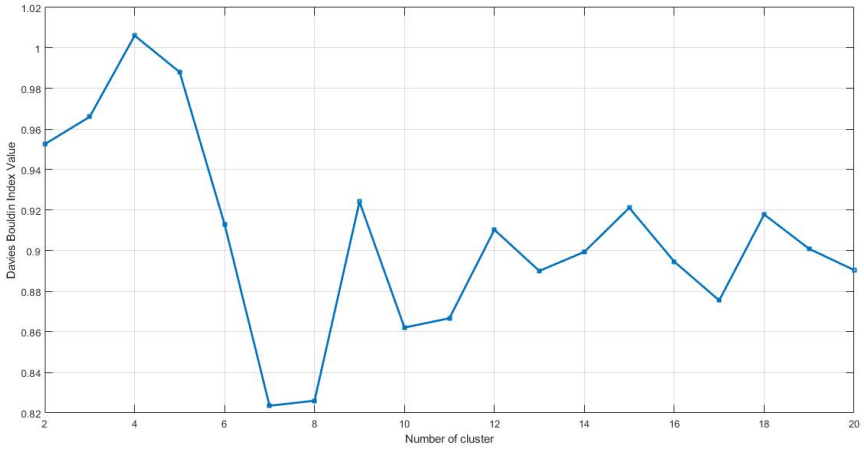
۴. تحلیل داده‌ها و یافته‌های پژوهش

در اجرای مدل تحلیل نتایج از نرم‌افزار متلب استفاده شد. برای مدل‌سازی بر اساس مشتری نیاز به ایجاد انباره داده‌ای از مشتریان است؛ بنابراین از روی داده‌های موجود در هفت جدول سه گروه داده در ارتباط با مشتریان تهیه شد. این سه گروه شامل اطلاعات RFM مربوط به دریافت کد ملّی کالا (ایرانکد)، RFM مربوط به دریافت کد بین‌المللی کالا (GS1) و اطلاعات RFM مربوط به تعداد اشتراک سالیانه که در جدول ۴، اطلاعاتی در مورد ویژگی‌های ارتباطی هر یک از گروه‌ها و مقادیر آن‌ها ارائه شده است.

جدول ۴. اطلاعات مربوط به پارامترهای RFM هر خدمت

نوع اطلاعات	ویژگی‌ها	ویژگی‌های ارتباطی استفاده‌شده	توضیح
اطلاعات تراکنشی (پارامترهای RFM) مربوط به دریافت ایرانکد	رتبه تازگی (RNC)	تاریخ ایجاد در جدول ایرانکد	عاملی است متناسب با آخرین تراکنش سرویس ایرانکد
	تعداد تراکنش (FNC)	تاریخ ایجاد جدول ایرانکد	تعداد تراکنش مالی سرویس ایرانکد در طی ۶ سال
	عامل مالی (MNC)	تاریخ خدمت، هزینه خدمت، مبالغ پرداختی عضو، تاریخ پرداخت	مجموع مقدار پول پرداختی از دریافت ایرانکد در طی ۶ سال
اطلاعات تراکنشی (پارامترهای RFM) مربوط به کد بین‌المللی (GS1)	رتبه تازگی (RGS1)	تاریخ ایجاد در جدول GS1	عاملی است متناسب با آخرین تراکنش سرویس GS1
	تعداد تراکنش (FGS1)	تاریخ ایجاد جدول GS1	تعداد تراکنش مالی سرویس GS1 در طی ۶ سال
	عامل مالی (MGS1)	تاریخ خدمت، هزینه خدمت، مبالغ پرداختی عضو، تاریخ پرداخت	مجموع مقدار پول پرداختی از دریافت کد GS1 در طی ۶ سال
اطلاعات تراکنشی (پارامترهای RFM) مربوط به عضویت/ تمدید اشتراک	رتبه تازگی (RMT)	تاریخ تراکنش	عاملی است متناسب با آخرین تراکنش تمدید اشتراک
	تعداد تراکنش (FMT)	تاریخ تراکنش	تعداد تراکنش مالی تمدید اشتراک در طی ۶ سال
	عامل مالی (MMT)	تاریخ خدمت، هزینه خدمت، مبالغ پرداختی عضو، تاریخ پرداخت	مجموع مقدار پول پرداختی از تمدید اشتراک در طی ۶ سال

در مرحله بعد مقادیر RFM هر یک از خدمات را نرمال‌سازی می‌شود؛ سپس با استفاده از شاخص دیویس بولدین، تعداد خوشه بهینه به دست می‌آید. شکل ۲، بهترین تعداد خوشه را که ۷ است، نشان می‌دهد.



شکل ۲. نمودار شاخص دیویس بولدین

پس از تعیین تعداد خوشه‌های بهینه، برای خوشه‌بندی مشتریان به‌منظور تفکیک آن‌ها در گروه‌های همگن بر اساس ارزش شاخص‌های مدل، از روش خوشه‌بندی K-Means استفاده شده است. مقدار متوسط پارامترهای R، F، M هر خدمت برای هر یک از خوشه‌ها در جدول ۵، مشاهده می‌شود.

جدول ۵. نتایج خوشه‌بندی اعضا بر اساس پارامترهای RFM برای هر خدمت

خوشه	خدمات ایرانکد			خدمات GS1			خدمات عضویت/تمدید اشتراک		
	FNC	RNC	MNC	FGS1	RGS1	MGS1	FMT	RMT	MMT
۱	۰/۰۰۰۲	۰/۰۰۲۵	۰/۰۰۰۱	E-۸/۵۸۰۶	۰۵E-۴/۷۲	۰E+۰/۰۰	۰/۱۵۹۵	۰/۸۸۱۴	۰/۱۵۹۶
۲	۰/۰۰۰۴	۰/۰۵۷۶	۰/۰۰۰۷	E-۱/۰۶۰۴	۰۳E-۵/۴۴	۰	۰/۱۰۱۰	۰/۰۸۴۹	۰/۰۶۶۰
۳	۰/۰۰۲۲	۰/۸۵۹۳	۰/۰۱۰۶	۰/۰۰۶۲	۰/۸۶	۰/۰۱۳۷	۰/۴۹۹۲	۰/۸۷۴۳	۰/۳۴۴۵
۴	۰/۰۰۰۴	۰/۷۷۹	۰/۰۰۲۱	E-۳/۵۴۰۵	۰۳E-۷/۱۵	۰۵E-۴/۸۴	۰/۱۹۰۸	۰/۸۱۱۳	۰/۱۹۴۴
۵	۰/۰۰۱۵	۰/۲۴۶۲	۰/۰۰۳۹	۰/۰۰۰۴	۰/۰۳۸۱	۰/۰۰۰۱	۰/۲۷۳	۰/۳۴۰۷	۰/۱۷۵۵
۶	۰/۰۰۲۶	۰/۷۳۹۶	۰/۰۰۸۳	۰/۰۰۰۹	۰/۰۶۵۳	۰۴E-۷/۵	۰/۷۶۵	۰/۸۷۳۲	۰/۴۴۹۶
۷	۰/۰۰۰۹	۰/۱۵۲۶	۰/۰۰۱۹	E-۴/۷۵۰۴	۰/۰۳۲۵	۰۴E-۱/۱	۰/۷۴۷۷	۰/۸۳۲۸	۰/۴۰۳۹

برای تعیین وزن‌های مربوط به شاخص‌های مدل RFM از فرد خبره در مرکز کمک گرفته شده است.

وزن هر یک از پارامترهای R، F، M، به‌ترتیب به‌صورت WR، WF، WM نمایش داده می‌شود که مجموع نسبی آن‌ها، برابر با یک است که در رابطه ۴، نشان داده شده است.

$$W_R + W_F + W_M = 1$$ رابطه (۴)

وزن هر یک از شاخص‌ها برای هر خدمت در جدول ۶، ارائه شده است.

جدول ۶. وزن‌های پارامترهای RFM هر خدمت

وزن	FNC	RNC	MNC	FGS1	RGS1	MGS1	FMT	RMT	MMT
w	۰/۵	۰/۲	۰/۳	۰/۵	۰/۲	۰/۳	۰/۵	۰/۲	۰/۳

امتیاز CLV هر یک از خوشه‌ها را بر اساس تحلیل RFM وزن‌دارشده به‌صورت رابطه ۵، محاسبه می‌شود:

$$CLV_{ci} = (NRNC_{ci} \times WRNC_{ci}) + (NFNC_{ci} \times WFNC_{ci}) + (NMNC_{ci} \times WMNC_{ci}) + (NFGS1_{ci} \times WFGS1_{ci}) + (NRGS1_{ci} \times WRGS1_{ci}) + (NMGS1_{ci} \times WMGS1_{ci}) + (NFMT_{ci} \times WFMT_{ci}) + (NRMT_{ci} \times WRMT_{ci}) + (NMMT_{ci} \times WMMT_{ci})$$ رابطه (۵)

در رابطه ۵، NRNC مقدار نرمال‌شده پارامتر RNC در خوشه ci و WRNC وزن پارامتر RNC در خوشه ci است. سایر پارامترها نیز به همین صورت تعریف می‌شوند. جدول ۷، مقدار CLV به‌دست‌آمده و رتبه هر خوشه را نشان می‌دهد.

جدول ۷. نتایج به‌دست‌آمده محاسبه CLV و رتبه‌بندی اعضا

خوشه	CLV	درصد اعضا هر در خوشه	رتبه	نامگذاری خوشه	وضعیت متوسط ارزش شاخص‌ها (FRMNC FRMGS1 FRMMT)
۱	۰/۳۰۴۵	٪۵/۴	F	اعضای جدید کم‌ارزش	↓↓↓ ↓↓↓ ↓↑↓
۲	۰/۱۰۰۴	٪۳۰/۷	G	اعضای کم‌ارزش	↓↓↓ ↓↓↓ ↓↓↓
۳	۰/۸۸۳۳	٪۱۴/۴	A	اعضای وفادار	↑↑↑ ↑↑↑ ↑↑↑
۴	۰/۴۷۴۱	٪۷/۴	D	اعضای جدید بالقوه	↓↑↓ ↓↓↓ ↓↑↓
۵	۰/۳۱۶۴	٪۲۸/۷	E	اعضای متوسط	↑↓↓ ↓↓↓ ↓↓↓
۶	۰/۸۵۷۵	٪۴/۴	B	اعضای ارزشمند	↑↑↑ ↓↓↓ ↑↑↑
۷	۰/۷	٪۹	C	اعضای خوب	↓↓↓ ↓↓↓ ↑↑↑

ابتدا در قسمت وضعیت شاخص‌ها، متوسط ارزش هر یک از شاخص‌های مدل RFM هر خدمت در هر خوشه با متوسط ارزش این شاخص‌ها در کل داده‌ها که در جدول ۵، نمایش داده شده است، مقایسه می‌شود که در این مقایسه مشخص شده است که متوسط ارزش هر یک از شاخص‌های مدل پیشنهادی در هر خوشه نسبت به متوسط ارزش این شاخص‌ها در کل داده‌ها در چه وضعیتی است. برای هر یک از شاخص‌ها در وضعیتی که متوسط ارزش شاخص در یک خوشه بیشتر از متوسط ارزش آن شاخص در کل داده باشد، این وضعیت با علامت (↑) (وضعیت مطلوب) و درصورتی‌که متوسط شاخص در یک خوشه کمتر از متوسط ارزش آن در کل داده باشد، این وضعیت با علامت (↓) (وضعیت نامطلوب) نشان داده می‌شود.

مدل‌های پیش‌بینی. هدف از اجرای این روش‌ها در مدل پیشنهادی، پیش‌بینی کلاس یا طبقه مشتریان بر مبنای خوشه‌بندی است. در این مرحله به‌منظور طبقه‌بندی مشتریان داده‌های موجود به دو مجموعه آموزش با تعداد ۵۹۵۰۸ عضو و آزمایش با ۱۴۸۷۷ عضو تقسیم شدند. از مجموعه داده‌های آموزش برای یادگیری و مدل‌سازی و از مجموعه آزمایش برای ارزیابی مدل و دقت پیش‌بینی استفاده شد. در این قسمت شبکه‌های عصبی و درخت تصمیم که زیرمجموعه‌ای از روش‌های هوش مصنوعی به‌شمار می‌آیند و همچنین رگرسیون لجستیک استفاده شده است. از میان انواع الگوریتم‌های درختان تصمیم از الگوریتم C4.5 برای ساخت مدل‌های طبقه‌بندی بهره‌گیری می‌شود.

نتایج دقت پیش‌بینی. مهم‌ترین معیار برای تعیین کارایی تکنیک دسته‌بندی معیار صحت[1] است. این معیار دقت کل یک دسته‌بندی را محاسبه می‌کند و نشان‌دهنده این حقیقت است که دسته‌بندی طراحی‌شده چند درصد از کل مجموعه رکوردهای آزمایشی را به‌درستی دسته‌بندی کرده است.

جدول ۸. متوسط دقت، یادآوری و صحت مدل‌ها

الگوریتم طبقه‌بندی	دقت[3]	بازخوانی[2]	صحت
شبکه‌عصبی	٪۹۹/۵۶	٪۹۹/۲۳	٪۹۹/۲۷
درخت تصمیم	٪۹۹/۳۸	٪۹۹/۱۵	٪۹۹/۱۲
رگرسیون لجستیک	٪۹۷/۹۲	٪۹۶/۹۲	٪۹۷/۲۲

1. Accuracy
2. Recall
3. Precision

نتایج جدول ۸، نشان می‌دهد که شبکه عصبی طراحی‌شده برای طبقه‌بندی اعضا، ۹۹/۵۶ درصد داده‌های آزمایشی را به‌درستی تشخیص داده است و با توجه به بالابودن درصد سه معیار دقت، بازخوانی و صحت روش شبکه عصبی می‌توان نتیجه گرفت که این روش نسبت به دو روش دیگر کارآمدتر است.

یافته‌ها. بر اساس یافته‌های پژوهش می‌توان به سؤال‌های مطرح‌شده در مقدمه پاسخ داد. می‌توان به سؤال یک که چگونگی خوشه‌بندی را مطرح کرده است، این‌گونه پاسخ داد که اعضای «مرکز ملّی شماره‌گذاری کالا و خدمات ایران» توسط تکنیک خوشه‌بندی K-Means و مدل RFM و شاخص دیویس بولدین به هفت خوشه تفکیک شده‌اند. در پاسخ به سؤال دو مبنی بر چگونگی محاسبه CLV اعضای مرکز می‌توان گفت که میزان ارزش طول عمر هر خوشه (CLV) توسط مدل وزنی RFM محاسبه می‌شود. پاسخ داده‌شده به سؤال سوم، یعنی چگونگی سنجش ارزش اعضا بر اساس CLV، به این صورت است که رتبه‌بندی و تحلیل هفت خوشه بر اساس ارزش CLV هر یک از آن‌ها انجام شده است؛ به‌طوری‌که ارزش CLV بیشتر نشانه رتبه بالاتر خوشه است. با توجه به نتایج این پژوهش و یکی از مسائل مطرح‌شده در سازمان مبنی بر محدودیت تعداد کدهای شرکتی و اینکه کدام اعضای غیرفعال را حذف و کدهای آن‌ها آزاد تا به اعضای جدید اختصاص داده شود، می‌توان پاسخ داد: به‌طور قطع اعضایی که در خوشه دو قرار گرفته و CLV آن‌ها پایین‌تر است، می‌توانند در زمره اعضای حذفی قرار گیرند. در پاسخ به سؤال آخر که چگونگی ارائه مدل پیش‌بینی را مطرح کرده است، می‌توان این‌گونه پاسخ داد که داده‌ها توسط سه الگوریتم طبقه‌بندی (درخت تصمیم C4.5، شبکه عصبی چندلایه و رگرسیون لجستیک) مدل‌سازی و با توجه به نتیجه ارزیابی عملکرد هر الگوریتم، شبکه عصبی چندلایه با دقت بالاتر (۹۹/۵۶ درصد) نسبت به دو الگوریتم دیگر به‌عنوان الگوریتم بهتر مشخص شد. با توجه به پژوهش‌های پیشین می‌توان گفت: «تکنیک‌های داده‌کاوی می‌تواند کمک بسزایی به شناسایی مشتریان و رتبه‌بندی آن‌ها و به‌طورکلی در مدیریت ارتباط با مشتریان کند».

۵. نتیجه‌گیری و پیشنهادها

تمامی سازمان‌های مشتری‌محور برای استفاده بهینه از منابع و ارائه خدمات مناسب به هر یک از مشتریان خود نیازمند شناسایی و تحلیل رفتار آن‌ها است. تکنیک‌های داده‌کاوی می‌تواند با بخش‌بندی مشتریان و تحلیل هر یک از بخش‌ها با شناسایی و تحلیل رفتار مشتریان به سازمان‌ها کمک بسزایی کند. در این پژوهش از مدل مناسب RFM برای خوشه‌بندی و محاسبه ارزش هر یک از مشتریان و تحلیل آن‌ها استفاده و مدل پیش‌بینی شبکه‌های عصبی به‌عنوان بهترین مدل پیش‌بینی انتخاب شد. با توجه به نتایج می‌توان پیشنهادهایی را برای هر یک از

بخش‌های مشتریان ارائه داد. به صورتی که برای اعضای وفادار (طلایی) که دارای جذابیت بالایی هستند، می‌توان با اعطای هدایایی به مناسبت‌های مختلف به آن‌ها، وفاداری آنان به مرکز را قطعیت بخشید. داشتن رابطه مستمر و همیشگی با این دسته از مشتریان از طریق تشکیل باشگاه مشتریان و ارتباطات پیامکی و همچنین بستن قردادهای طولانی‌مدت به‌منظور شراکت استراتژیک تا حد زیادی می‌تواند بقای این مشتریان در درازمدت را برای مرکز تضمین کند و سودآوری مرکز را ارتقا بخشد. اعضای ارزشمند (پلاتینی) مانند دسته نخست دارای جذابیت بالا برای مرکز هستند؛ اما ارزش طول عمر کمتری نسبت به دسته نخست دارند. اعمال سیاست‌های مشابه با دسته نخست برای این دسته از مشتریان نیز پیشنهاد می‌شود. اعضای خوب (نقره‌ای)، از نظر تمدید قرارداد وضعیت نسبتاً خوبی دارند؛ ولی در خرید از کدهای GS1 و ایرانکد وضعیت مطلوبی ندارند. برای این دسته از مشتریان برگزاری دوره‌های آموزشی یا استفاده از تبلیغات بصری، نظیر چاپ بروشور و غیره و همچنین توصیف قابلیت‌ها و مزایای حاصل از سیستم‌های کدینگ ملّی و بین‌المللی پیشنهاد می‌شود. اعضای جدید بالقوه (برنزی) جزو مشتریانی هستند که اخیراً محصولات مرکز را خریداری کرده‌اند. با توجه به جدیدالورودبودن آن‌ها می‌توان امیدوار بود با برنامه‌ریزی کارا در دسته مشتریان سودآور قرار گیرند؛ بنابراین معرفی انواع محصولات مرکز به آن‌ها از طریق برگزاری همایش، تبلیغات تلفنی، پیامکی و غیره می‌تواند تا حد زیادی به ارتقا سطح درآمد مرکز کمک کند. اعضای متوسط (برنجی)، غالباً مشتریان دارای ارزش بالا برای مرکز نیستند. با تحلیل ویژگی‌های این مشتریان مشخص می‌شود که عمده این مشتریان مربوط به دوره‌ای از کار مرکز هستند که خرید کدها ایرانکد دارای الزام قانونی بوده است. درآمدزایی از این مشتریان بیشتر از آنکه به سیاست‌های بازاریابی وابسته باشد به ارتقا عملکرد فنی شرکت در راستای ایجاد ارزش‌افزوده برای استفاده از سیستم‌های کدینگ کالا و خدمات و نیز افزایش تنوع محصولات مرکز است. اعضای جدید کم‌ارزش (آهنی و سربی) جزو کم‌ارزش‌ترین مشتری‌ها برای مرکز هستند و سرمایه‌گذاری بر روی این مشتریان توصیه نمی‌شود. با توجه به محدودیت‌های موجود در زمینه صدور کدهای ایرانکد و GS1 پیشنهاد می‌شود کدهای صادرشده برای این مشتریان حذف و برای مشتریان جدید استفاده شود.

محدودیت‌های این این پژوهش عبارت‌اند از: ۱. دسترسی‌نداشتن به بازه زمانی تراکنش‌های قبل از سال ۱۳۹۰. با توجه به آنکه در سال ۱۳۹۲ ایرانکد از حالت اجباری خارج‌شده می‌توان مقایسه‌ای بر اساس بازه‌های قبل از سال ۱۳۹۲ و بعد از آن انجام داد و به نتایج دقیق‌تری دست‌یافت؛ ۲. دسترسی‌نداشتن به برخی از جدول ها برای به‌دست‌آوردن وضعیت اعضا و یا حقیقی/ حقوقی بودن آن‌ها، اطلاعات خارج سازمان از جمله میزان سود اعتبار، نرخ مالیات هر یک از اعضا و همچنین محدودیت دسترسی به مقاله‌های به‌روز و مرتبط با موضوع این پژوهش.

پیشنهادهایی را برای پژوهش‌های آتی می‌توان بیان کرد از جمله:

پیشنهادهای نظری:

ـ کارکرد سایر روش‌های خوشه‌بندی نظیر نقشه‌های خودسازمان‌ده، خوشه‌بندی سلسله‌مراتبی و یا روش‌های ترکیبی حاصل از الگوریتم‌های فراابتکاری و الگوریتم K-means برای خوشه‌بندی مشتریان اجرا شده و نتایج آن‌ها با یکدیگر مقایسه شود؛

ـ با توجه به توسعه و مزیت‌های موجود در روش‌های فازی پیشنهاد می‌شود از روش‌های مبتنی بر مدل‌های فازی برای این منظور استفاده شده و نتایج آن‌ها با رویکردهای قطعی مقایسه شود.

پیشنهادهای کاربردی:

ـ شاخص‌های دیگری نظیر حوزه کاری مشتری، حقیقی‌بودن یا حقوقی‌بودن آن‌ها و غیره در نظر گرفته شود؛

ـ با درنظرگرفتن تأثیر متغیرهای خارجی، مانند نرخ تورم، نرخ مالیات، اعتبار مالی و غیره در پیش‌بینی رفتار مشتریان به پیش‌بینی‌های دقیق‌تری از رفتار مشتریان دست یافت؛

ـ به دلیل لغو اجباری‌بودن ایرانکد در مهرماه سال ۱۳۹۲ می‌توان عملکرد اعضا را قبل و بعد از این رخداد بررسی کرد؛

ـ می‌توان از نتایج این پژوهش برای ایجاد یک سیستم تصمیم‌یار (DSS)[1] به‌منظور کمک به مدیران «مرکز ملّی شماره‌گذاری کالا و خدمات» استفاده کرد.

1. Decision Suport System

منابع

1. Anitha, P., & Patil, M. M. (2019). RFM model for Customer Purchase Behavior using K-Means Algorithm. Journal of King Saud University-Computer and Information Sciences, 1319-1578.
2. Carolyn, F. C., & Karen, N. K. (2002). From prisoners to apostles: a typology of repeat buyers and loyal customers in service businesses. *Journal of services Marketing, 16(4),* 322-341.
3. Çavdar, A. B., & Ferhatosmanoğlu, N. (2018). Airline customer lifetime value estimation using data analytics supported by social network information. *Journal of Air Transport Management, 67,* 19-33.
4. Cheng, C.-H., & Chen, Y.-S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications, 36(3),* 4176-4184.
5. Cheng, C.-J., Chiu, S., Cheng, C.-B., & Wu, J.-Y. (2012). Customer lifetime value prediction by a Markov chain based data mining model: Application to an auto repair and maintenance company in Taiwan. *Scientia Iranica, 19(3),* 849-855.
6. Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2018). RFM ranking–An effective approach to customer segmentation. *Journal of King Saud University-Computer and Information Sciences.*
7. Daneshvar, A., Homayounfar, M., FarahmandNezhad, A. (2020), Development of an intelligent multi-criteria clustering method based on Promethee. *Industrial Management Perspective, 36,* 41-6. (In Persian)
8. Dimitriadis, S., & Stevens, E. (2008). Integrated customer relationship management for service activities: an internal/external gap model. *Managing Service Quality: An International Journal, 18(5),* 496-511.
9. Dursun, A., & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives, 18,* 153-160.
10. Farokhi, z. (2013). Segmentation of bankcard holders based on LRFM model using data mining techniques. *Brand managemen Journal. (*In Persian)
11. Haenlein, M., Kaplan, A. M., & Beeser, A. J. (2007). A model to determine customer lifetime value in a retail banking context. *European Management Journal, 25(3),* 221-234.
12. Heldt, R., Silveira, C. S., & Luce, F. B. (2019). Predicting customer value per product: From RFM to RFM/P. Journal of Business Research, 0148-2963.
13. Hiziroglu, A., & Sengul, S. (2012). Investigating two customer lifetime value models from segmentation perspective. *Procedia-Social and Behavioral Sciences, 62,* 766-774.
14. Iranshahi, M. (2015), Investigating the Necessity and effect of CLV to using CRM in DANA Insurance Co. (In Persian)
15. Jarahi, M., Ardakani, S., & Zareiyan, M. (2009). Investigating the role of information technology in implementing CRM electronically (eCRM). *Quarterly Journal of Parks and Roshd Centers, 21.* (In Persian)
16. Khajvand, M., & Tarokh, M. J. (2011). Estimating customer future value of different customer segments based on adapted RFM model in retail banking context. *Procedia Computer Science, 3,* 1327-1332.

17. Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science, 3,* 57-63.

18. Khatami FiroozAbadi, M., TaghaviFard, M., Sadjadi, Kh., Bamdad Soufi, J. (2018), Optimization through simulation to solve the problem of multi-objective allocation of services to the bank's clustered customers. *Industrial Management Perspective, 30,* 85-110. (In Persian)

19. Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: the construct, research propositions, and managerial implications. The Journal of Marketing, *54(2),* 1-18.

20. Larose, D. T. (2006). *Data mining methods & models.* John Wiley & Sons.

21. Li, D. C., Dai, W. L., & Tseng, W. T. (2011). A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. *Expert Systems with Applications, 38(6),* 7186-7191.

22. Monalisa, S., Nadya, P., & Novita, R. (2019). Analysis for Customer Lifetime Value Categorization with RFM Model. *Procedia Computer Science, 161,* 834-840.

23. Moslehi, N., Kafashpour, A., & *Naji Azimi, Z. (2014). Customer segmentation base on their CLV using LRFM model, Management Researches Journal, 7(25), 2014119-14*0. (In Persian)

24. Motameni, A., Rezaei, M., & Ehghaghi, M. (2013). Designing a demand perdiction model in the ceramic and tile industry. *Industrial Management Perspective, 9,* 159-176. (In Persian)

25. Qadadeh, W., & Abdallah, S. (2018). Customers Segmentation in the Insurance Company (TIC) Dataset. *Procedia computer science, 144,* 277-290.

26. Romano Jr, N. C., & Fjermestad, J. (2001). Electronic commerce customer relationship management: An assessment of research. *International Journal of Electronic Commerce, 6(2),* 61-113.

27. Tabeli, H., & PourJafari, M. (2011). Role of information technology in the CRM development, *Second National Conference on Information and Communication Technology* (In Persian)

28. Tamaddoni, A., Stakhovych, S., & Ewing, M. (2017). The impact of personalised incentives on the profitability of customer retention campaigns. *Journal of Marketing Management, 33(5-6),* 327-34.

29. Tarokh, M. J., & EsmaeiliGookeh, M. (2019). Modeling patient's value using a stochastic approach: An empirical study in the medical industry. *Computer methods and programs in biomedicine, 176,* 51-59.

30. Taslimi, H., Aghazadeh Hashem, M., Esfidani, M., & Karami, M. (2003). Critique of marketing philosophies. *Management Knowledge Quarterly, 16(61-60),* 3-21. (In Persian)

31. Wang, Y., Po Lo, H., Chi, R., & Yang, Y. (2004). An integrated framework for customer value and customer-relationship-management performance: a customer-based perspective from China. *Managing Service Quality: An International Journal, 14(2/3),* 169-182.