

## Branch Client Behavior Analysis Using RFM Method

**Fatemeh Rahimi** 

MSc., Industrial Engineering, Optimization  
Orientation, Faculty of Engineering, Kharazmi  
University, Tehran, Iran.

**Mohammad Vahid  
Sebt\*** 

Assistant Professor, Faculty of Engineering,  
Kharazmi University, Tehran, Iran.

**Nasim Ghanbar  
Tehrani** 

Assistant Professor, Faculty of Engineering,  
Kharazmi University, Tehran, Iran.

### Abstract

In today's competitive world, applying new techniques to business development has a great impact. The restaurant industry is no exception. Therefore, in this research, using new methods of knowledge discovery and data mining, customer data of chain restaurant is investigated. The purpose of this study was to explore customer behavior patterns using data mining methods. In this study, one million and five hundred thousand customer records were reviewed in five branches of a chain restaurant and two stages of clustering modeling using RFM method and then classification modeling were performed on the data and the behavior rules chain restaurant customers were extracted. The results of this study have helped to identify the loyal and profitable customers of the chain restaurant which has led to the improvement of the profitability of the chain restaurant. One of the innovations of this research has been the communication between clustering and classification results.

**Keywords:** Data Mining, Customer Relationship Management, K-Means Algorithm, RFM, Loyalty.


\* Corresponding Author: Vhd\_sebt@yahoo.com

**How to Cite:** Sebt, M, V., Rahimi, F., Ghanbar Tehrani, N., (2020). Branch Client Behavior Analysis Using RFM Method, *Journal of Business Intelligence Management Studies*, (36), 209-242.




## تحلیل الگوی رفتاری مشتریان شعب به روش خوشه‌بندی و دسته‌بندی با استفاده از روش RFM


کارشناسی ارشد مهندسی صنایع، گرایش بهینه‌سازی، دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران.

فاطمه رحیمی 

استادیار دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران.

محمد وحید سبط  \*

استادیار دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران.

نسیم غنبر تهرانی 

### چکیده

در دنیای پررقابت امروز، به‌کارگیری تکنیک‌های جدید در پیشرفت کسب‌وکار تأثیر به‌سزایی دارد. صنعت رستوران نیز از این قاعده مستثنا نیست. از این‌رو، در این پژوهش، با استفاده از روش‌های نوین کشف دانش و داده‌کاوی به بررسی داده‌های مشتریان رستوران زنجیره‌ای پرداخته شده است. هدف از این تحقیق، کشف الگوهای رفتاری مشتریان با استفاده از روش‌های داده‌کاوی بوده است. در این پژوهش، تعداد یک‌میلیون و پانصد هزار رکورد از سوابق مشتریان در ۵ شعبه یک رستوران زنجیره‌ای مورد بررسی قرار گرفته‌اند و دو مرحله مدل‌سازی خوشه‌بندی با استفاده از روش RFM و سپس مدل‌سازی دسته‌بندی بر روی داده‌ها اجرا گردید و قواعد رفتاری مشتریان رستوران زنجیره‌ای استخراج گردید. نتایج به‌دست‌آمده از این تحقیق، به شناخت مشتریان وفادار و سودآور رستوران زنجیره‌ای کمک کرده است که نهایتاً به بهبود سودآوری رستوران زنجیره‌ای منجر گردیده است. برقراری ارتباط بین نتایج به‌دست‌آمده از خوشه‌بندی و دسته‌بندی از جمله نوآوری‌های این تحقیق بوده است.

**کلیدواژه‌ها:** داده‌کاوی، مدیریت ارتباط با مشتری، الگوریتم K-Means، روش RFM، وفاداری.

## مقدمه

امروزه در شرکت‌های آینده‌نگر، پایه‌ریزی یک رابطه متداوم و استراتژیک با مشتریان، به‌منظور بقای کسب‌وکار و همچنین رشد و توسعه کسب‌وکار در اولویت قرار داده شده است. شناخت تمام مشتریان و همچنین شناخت نیازهای آن‌ها کمک بزرگی در راستای تحقق این هدف است.

در این نوع شرکت‌ها با توجه به تنوع زیاد مشتریان، تنوع کسب‌وکار، تنوع در پیچیدگی خدمات و نیازهای مشتریان و تنوع در دسترسی به اطلاعات آن‌ها، استفاده از یک روش مدون که به کشف روابط و دانش‌های پنهان کمک کند، احساس می‌شود، به همین دلیل، استفاده از داده‌کاوی برای طبقه‌بندی و کشف و تولید اطلاعات مناسب از مشتریان در میان حجم انبوهی از داده‌های شرکت‌ها و سازمان‌ها حائز اهمیت است. یکی از صنعت‌هایی که اخیراً با تغییرات زیادی روبه‌رو شده است، صنعت رستوران است با توجه به اینکه مشتریان بیش‌ازپیش به سلامت خود اهمیت می‌دهند و رستوران‌های بسیاری در دسترس هستند، مدیران رستوران‌ها بایستی استراتژی‌های بازاریابی، جذب و حفظ مشتریان خود را تغییر دهند. شناخت مشتریان و همچنین شناخت نیازهای آن‌ها، عمل مؤثری در کسب برتری در ارائه خدمات به مشتریان است. از این‌رو، مدیران شرکت‌ها درصدد آن هستند که مشتریان خود را اولویت‌بندی کرده و کانون توجه خود را بر مشتریان کلیدی متمرکز کنند.

خوشه‌بندی<sup>۱</sup> یکی از مباحث مطرح در حوزه مدیریت ارتباط با مشتری است که جمعیت زیادی از مشتریان را به بخش‌های مختلف تقسیم‌بندی می‌کند، به‌طوری‌که مشتریان موجود در هر بخش به یکدیگر شبیه و مشتریان بخش‌های مختلف با یکدیگر متفاوت باشند و دیدگاهی کلی از تمام بانک اطلاعاتی مشتریان را ارائه دهد. دانش کشف‌شده از رفتار مشتریان، به صاحبان کسب‌وکار امکان پایه‌گذاری استراتژی‌های

مختلف بازاریابی، جذب و حفظ مشتریان هر بخش را می‌دهد که در نهایت به بهبود سودآوری سازمان منجر می‌گردد.

### مبانی نظری پژوهش

یکی از مدل‌های مطرح شده در تحلیل ارزش مشتری، مدل RFM است که هاگس آن را برای نخستین بار در سال ۱۹۹۴ ارائه کرده است. این مدل با استفاده از سه متغیر تازگی، تکرار و ارزش پولی، به بیان تفاوت میان مشتریان می‌پردازد. مدل RFM یک مدل مبتنی بر رفتار است که برای آنالیز رفتار یک مشتری و سپس پیش‌بینی کردن مشتری بر اساس رفتار او در بانک اطلاعات استفاده می‌شود.

پارامترهای این مدل به شرح زیر هستند:

تازگی آخرین خرید: مدت زمان بین آخرین تعامل تجاری باحال را نشان می‌دهد. هر چه این مدت کمتر باشد، R بیشتر است.

تکرار خرید: تعداد تراکنش‌ها را در یک بازه زمانی مشخص نشان می‌دهد، هر چه تعداد تکرار بیشتر باشد، F بزرگ‌تر است.

ارزش پولی خرید: ارزش پولی تعاملات را در بازی‌های خاص نشان می‌دهد. هر چه ارزش پولی بیشتر باشد، M بزرگ‌تر است.

در بین فاکتورهای RFM، فاکتور تأخر اغلب مهم‌ترین متغیر است، با این حال بر اساس مطالعات انجام شده گذشته، متغیرهای RFM در واقع Firm – Specific هستند و بر اساس طبیعت محصولات شرکت اهمیت آن‌ها فرق می‌کند.

### پیشینه پژوهش

بیروول کاودار و همکاران (۲۰۱۸) در مقاله‌ای تحت عنوان " تخمین ارزش طول عمر مشتریان ایرلاین با استفاده از آنالیز داده‌های شبکه‌های اجتماعی " با استفاده از رگرسیون چندخطی به خوشه‌بندی مشتریان، آنالیز روی گردانی و تخمین ارزش مشتری با استفاده از اطلاعات شبکه‌های اجتماعی پرداختند.

گاو و همکاران (۲۰۱۸) از طریق ارائه مدلی که روابط رقابتی را از بازبینی‌های آنلاین استخراج کرده و سه نوع شبکه روابط رقابتی را می‌سازد برای آنالیز ساختار بازار، شناسایی رقبا و تعیین نقاط قوت و ضعف رستوران استفاده کردند. این پژوهش در تحقیقی تحت عنوان "شناسایی رقبا از طریق کاویدن روابط مقایسه‌ای بازبینی‌های آنلاین در صنعت رستوران" صورت گرفته است که از طریق ایده کاوی نظرات مشتریان و گراف چندخطی، نقاط قوت و ضعف رستوران‌های مختلف تعیین شده است.

فروزنده (۱۳۹۶) در تحقیقی تحت عنوان "پیش‌بینی رویگردانی مشتریان حقیقی بانک ملت با استفاده از روش‌های جمعی مبتنی بر بگینگ و بوستینگ" با استفاده از روش‌های جمعی مبتنی بر بگینگ و بوستینگ به متوازن نمودن داده‌ها در الگوریتم‌های استاندارد طبقه‌بندی پرداخته است و با استفاده از روش‌های جمعی رویگردانی مشتریان را مورد بررسی قرار داده است.

دورسان و همکاران (۲۰۱۶) در تحقیقی تحت عنوان "استفاده از تکنیک‌های داده کاوی به منظور گروه‌بندی مشتریان سودمند هتل" به خوشه‌بندی مشتریان سودمند هتل با استفاده از مدل‌سازی RFM پرداخته است.

سعیدی (۱۳۹۵) در مقاله‌ای تحت عنوان "خوشه‌بندی مشتریان استفاده‌کننده از بانکداری الکترونیک باهدف طراحی برنامه بازاریابی" با استفاده از روش RFM و الگوریتم K-Means به خوشه‌بندی مشتریان پرداخته است.

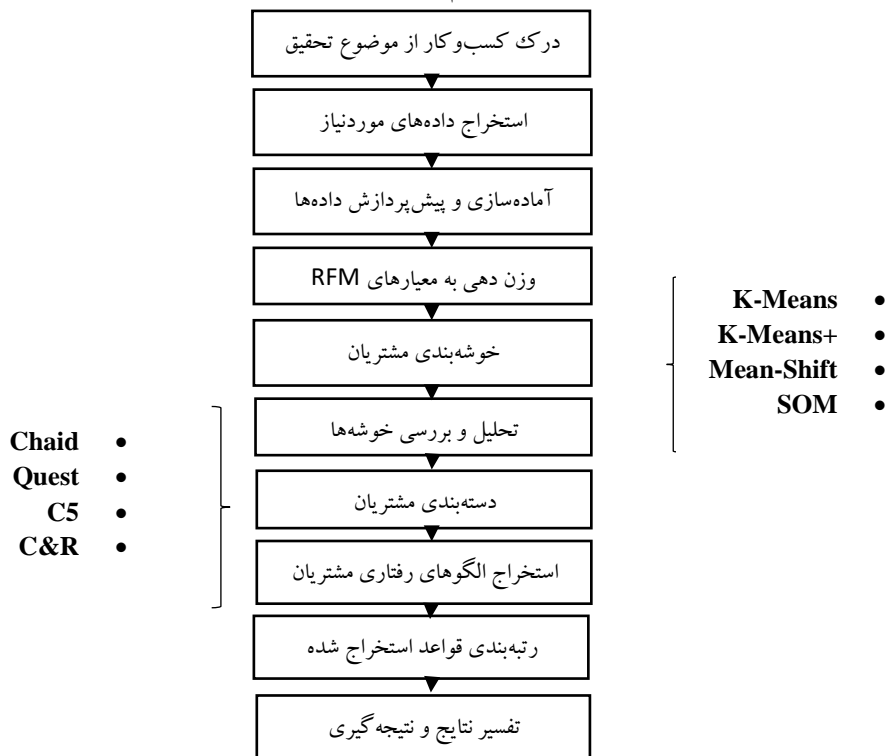
قاسم‌زاده (۱۳۹۵) در مقاله‌ای تحت عنوان "بخش‌بندی مشتریان اینترنت شبکه تلفن همراه و بررسی تغییرات رفتاری" با استفاده از الگوریتم‌های خوشه‌بندی به بخش‌بندی مشتریان و تحلیل رفتار مشتریان در بازه‌های مختلف پرداخته است.

ممتاز و همکاران (۱۳۹۳) در مقاله‌ای تحت عنوان "مدلی جدید برای ارزیابی رفتار مشتریان یک فست فود" به ارائه مدلی برای تعریف ارزش مشتریان بر مبنای مدل‌سازی RFM و خوشه‌بندی مشتریان یک رستوران فست فود پرداخت. هدف از این تحقیق خوشه‌بندی مشتریان رستوران، شناخت مشتریان وفادار و تدوین استراتژی‌هایی به منظور جذب و ابقاء هر خوشه از مشتریان بوده است.

در این پژوهش، ابتدا با استفاده از چهار الگوریتم K-Means، K-Means+، Mean-Shift و SOM خوشه‌بندی مشتریان یک رستوران زنجیره‌ای صورت گرفته است و در ادامه از طریق تلفیق نتایج حاصله از خوشه‌بندی با اطلاعات در دسترس از مشتریان نظیر محل زندگی و اطلاعات مرتبط با سفارش‌های آن‌ها، دسته‌بندی<sup>۱</sup> مشتریان و استخراج قواعد رفتاری انجام شده است.

### مدل پژوهش

در شکل ۱ مدل مفهومی اجرای تحقیق ترسیم شده است:



شکل ۱. مدل تحقیق

## روش‌شناسی پژوهش

پژوهش پیش رو در دسته تحقیقات عملی قرار گرفته است. در این پژوهش، هدف اصلی ارائه یا آزمون نظریه‌ها نیست، بلکه توسعه محصول (کالا یا خدمات) یا فرایندهای جدید است. نتایج حاصل از این پژوهش برای حل مسائل اجرایی و واقعی به کار گرفته شده است. این تحقیق بر اساس متدولوژی CRISP-DM<sup>۱</sup> انجام شده است.

در ابتدا داده موردنیاز جمع‌آوری گردیده است. در مرحله بعد، داده‌های جمع‌آوری شده وارد فاز پاک‌سازی داده‌ها شده‌اند. بخشی از پاک‌سازی داده‌ها از طریق فرمول نویسی در نرم‌افزار اکسل<sup>۲</sup> و بخشی از طریق کد نویسی به زبان پایتون<sup>۳</sup> انجام شده است. از داده‌های پاک‌سازی شده برای مدل‌سازی بهره گرفته شده است. مشتریان رستوران بر اساس روش RFM و از طریق کد نویسی به زبان پایتون خوشه‌بندی شده‌اند. پس از خوشه‌بندی مشتریان، فاز دوم مدل‌سازی با هدف دسته‌بندی مشتریان به دو دسته مشتریان خوب (سودآور) و مشتریان بد (با سودآوری کم) بر روی داده‌ها پیاده‌سازی شده است که به استخراج تعدادی از قواعد رفتاری مشتریان منجر گردیده است.

## تحلیل داده‌ها

با توجه به اینکه برای انجام فرایند پژوهش، استاندارد CRISP-DM مورد استفاده قرار گرفته است، در ادامه ساختار اجرایی پژوهش بر اساس مراحل این استاندارد تشریح شده است:

- درک مسئله کسب‌وکار
- جمع‌آوری و درک داده‌ها
- پیش‌پردازش و آماده‌سازی داده‌ها
- مدل‌سازی<sup>۴</sup>

---

1. Cross Industry Standard Process for Data Mining

2. Excel

3. Python

4. Modeling

- ارزیابی نتایج
- به کارگیری مدل

### درک مسئله کسب و کار و تعیین هدف اصلی

امروزه به دلیل در دسترس بودن محصولات و خدمات، شرکت‌ها به طور فزاینده‌ای نیازمند پایه‌گذاری روابط قدرتمند با مشتریان خود هستند. این بازار رقابتی گسترده، در صنعت رستوران بیش از پیش قابل مشاهده است. از این رو، این صنعت نیاز به شناخت مشتریان خود و کشف الگوهای رفتاری مشتریان با استفاده از روش‌های به‌روز دارد.

دلیل اصلی انجام این تحقیق، کمک به سودآوری رستوران زنجیره‌ای از طریق شناسایی گروه‌های مختلف مشتریان این رستوران بوده است. صاحب کسب و کار درصدد پاسخگویی به این سؤال بوده است که چگونه مشتریان وفادار خود را شناسایی کرده تا استراتژی‌هایی به‌منظور افزایش سودآوری رستوران تدوین گردد.

برای اینکه بتوان به مشتریان خدمات بهتری ارائه داد، باید در اولین قدم آن‌ها را شناخت، با شناخت هر دسته از مشتریان، می‌توان متناسب با نیاز آن‌ها، خدمات شرکت را بهبود داد و به این صورت مشتریان را حفظ کرد. پس از تشخیص مشتریان، مشخصات آن‌ها، خواسته‌ها، نیازها و انتظارات هر طبقه از آن‌ها و سپس شناخت خدمات شرکت، سودآوری شرکت افزایش می‌یابد.

مشتریان این رستوران زنجیره‌ای به دو صورت تلفنی و حضور در سالن، سفارش‌های خود را ثبت می‌کنند. لازم به ذکر است فقط مشتریانی که سفارش‌های خود را به صورت تلفنی ثبت کرده‌اند، کد مشتری دریافت کرده‌اند؛ بنابراین فقط داده مربوط به تحویل سفارش‌های رستوران در سال‌های ۱۳۹۵، ۱۳۹۶ و ۱۳۹۷ ذخیره شده در بانک اطلاعاتی رستوران در این تحقیق مورد استفاده قرار گرفته است تا امکان رصد اطلاعات سوابق خرید مشتریان فراهم شود.



## جمع‌آوری و درک داده‌ها

داده‌های مشتریان رستوران زنجیره‌ای در قالب یک دیتاست با یک میلیون و پانصد هزار رکورد از سوابق مشتریان در پنج شعبه این رستوران، واقع در مناطق جردن، پاسداران، تهران‌پارس، سعادت‌آباد و اندرزگو که سفارش‌های خود را به صورت تلفنی به رستوران ارسال کرده‌اند، جمع‌آوری شده است.

این تحقیق بر روی سفارش‌های ثبت‌شده در بازه زمانی سال ۱۳۹۵ تا ۱۳۹۷ انجام شده است. داده‌ها شامل اطلاعات دموگرافیک مشتریان، اطلاعات تراکنشی ثبت سفارش‌ها (تاریخ سفارش دهی، مبلغ خرید، محصولات خریداری‌شده، تعداد سفارش‌ها و...) و اطلاعات شعب بوده است که بر اساس موضوع تحقیق و مدل RFM تعدادی از متغیرها جهت اجرای تحقیق انتخاب گردیدند.

## پیش‌پردازش<sup>۱</sup> و آماده‌سازی داده‌ها<sup>۲</sup>

مجموعه داده ارائه‌شده از مشتریان رستوران زنجیره‌ای شامل یک میلیون و پانصد هزار رکورد و ۱۳ ویژگی بوده است. فیلدهای اطلاعاتی این دیتاست، شامل اطلاعاتی از قبیل شماره فاکتور، شماره مشتری<sup>۳</sup>، نام مشتری<sup>۴</sup>، تاریخ سفارش دهی<sup>۵</sup>، شعبه رستوران<sup>۶</sup>، کد محصول سفارش داده شده<sup>۷</sup>، نام محصول سفارش داده شده<sup>۸</sup>، آدرس<sup>۹</sup>، تعداد<sup>۱۰</sup>، قیمت<sup>۱۱</sup>، واحد<sup>۱۲</sup>، تخفیف<sup>۱۳</sup> و توضیحات بوده است. در جدول شماره ۱ اطلاعات کامل‌تری از

- 
1. Pre-Processing
  2. Data Preparing
  3. Customer ID
  4. Customer Name
  5. Date
  6. Branch Name
  7. Product Code
  8. Product Name
  9. Address
  10. Quantity
  11. Price
  12. Unit of Measurement
  13. Discount

فیلدهای اطلاعاتی (ویژگی‌ها) اولیه و تغییرات صورت گرفته در جریان پاک‌سازی داده‌ها ارائه شده است:

جدول ۱ توضیحات فیلدهای اطلاعاتی اولیه دیتاست

ردیف	ویژگی	نوع داده	محدوده مقادیر	تغییرات
۱	شماره فاکتور	Int	۲۰۱۸۶۷-۲۲۳	ویژگی فاقد اهمیت که حذف شد.
۲	شماره مشتری	Int	۱۰۰۲-۲۳۵۵۵	بکار گرفته شد.
۳	نام مشتری	String		ویژگی فاقد اهمیت که حذف شد.
۴	تاریخ سفارش دهی	Date	از اول فروردین ماه سال ۱۳۹۵ - ششم آذرماه سال ۱۳۹۷	تفکیک شد و به ۳ فیلد اطلاعاتی "روز"، "ماه" و "سال" تبدیل شد.
۵	شعبه رستوران	String	اندرزگو	ویژگی "شماره شعبه" <sup>۴</sup> از این فیلد اطلاعاتی ایجاد شد.
			جردن	
			تهران پارس	
			سعادت آباد	
	پاسداران			
۶	کد محصول سفارش داده شده	Int		بکار گرفته شد.
۷	نام محصول سفارش داده شده	String		فیلد اطلاعاتی "گروه محصول" <sup>۵</sup> از این فیلد اطلاعاتی ایجاد شد.

1. Day
2. Month
3. Year
4. Branch Code
5. Product Category

ردیف	ویژگی	نوع داده	محدوده مقادیر	تغییرات
۸	تعداد	Int	۱:۴	فیلدهای اطلاعاتی تعداد کل محصولات خریداری شده هر مشتری در سال‌های ۱۳۹۵، ۱۳۹۶ و ۱۳۹۷ خلق شدند.
۹	قیمت	Float		۳ فیلد اطلاعاتی "مجموع مبالغ سفارش دهی مشتری در سال ۹۵"، "مجموع مبالغ سفارش دهی مشتری در سال ۹۶" و "مجموع مبالغ سفارش دهی مشتری در سال ۱۳۹۷" از این فیلد اطلاعاتی ایجاد شد.
۱۰	واحد	String	عدد	ویژگی فاقد اهمیت که حذف شد.
۱۱	تخفیف	Float		ویژگی فاقد اهمیت که حذف شد.
۱۲	آدرس	String		ویژگی‌های "نام منطقه" <sup>۴</sup> و "شماره منطقه" <sup>۵</sup> از این فیلد اطلاعاتی ایجاد شده است.
۱۳	توضیحات	String		ویژگی فاقد اهمیت که حذف شد.

در مرحله اول، ویژگی‌هایی از جمله "روز"، "ماه" و "سال" از فیلد اطلاعاتی "تاریخ سفارش دهی" ایجاد شدند که جزئیات تاریخ سفارش دهی هر مشتری از طریق این ۳ فیلد اطلاعاتی مشخص شده است. همچنین، فیلدهای اطلاعاتی مانند "تعداد دفعات مراجعه مشتری در سال ۱۳۹۵"<sup>۶</sup>، "تعداد دفعات مراجعه مشتری در سال ۱۳۹۶"<sup>۷</sup> و "تعداد دفعات

- 
1. Total Order Amount Per Customer 1395
  2. Total Order Amount Per Customer 1396
  3. Total Order Amount Per Customer
  4. Region Name
  5. Region ID
  6. Customer Frequency 1395
  7. Customer Frequency 1396

مراجعه مشتری در سال ۱۳۹۷<sup>۱</sup> بر اساس فیلدهای اطلاعاتی "سال" و "شماره مشتری" و از طریق فرمول نویسی در اکسل خلق شدند.

ویژگی "شماره شعبه" از ویژگی "نام شعبه" ایجاد شده است. بر اساس فیلد اطلاعاتی "آدرس" ویژگی‌های "نام منطقه" و سپس "شماره منطقه" ایجاد شده است. ویژگی‌های "تعداد کل محصولات خریداری شده هر مشتری در سال ۱۳۹۵"<sup>۲</sup>، "تعداد کل محصولات خریداری شده هر مشتری در سال ۱۳۹۶"<sup>۳</sup> و "تعداد کل محصولات خریداری شده هر مشتری در سال ۱۳۹۷"<sup>۴</sup> از ویژگی "تعداد" به وجود آمده است. محصولات رستوران در ۹ گروه "برگر"<sup>۵</sup>، "ساندویچ"<sup>۶</sup>، "سوخاری"<sup>۷</sup>، "پیتزا"<sup>۸</sup>، "پیش غذا"<sup>۹</sup>، "دسر"<sup>۱۰</sup>، "سالاد"<sup>۱۱</sup>، "نوشیدنی طبیعی"<sup>۱۲</sup> و "نوشیدنی گازدار"<sup>۱۳</sup> قرار داده شدند و گروه محصول انتخابی هر مشتری، بر اساس فیلد اطلاعاتی "نام محصول سفارش داده شده" در فیلد اطلاعاتی دیگری با نام "گروه محصول" ثبت شدند. همچنین فیلدهای اطلاعاتی باینری با عناوین مشابه گروه محصول نیز ایجاد شده است که در صورت سفارش مشتری از هر یک از این گروه‌های محصول، مقدار ۱ در این فیلد اطلاعاتی نمایش داده شده است.

از ویژگی‌های "تعداد" و "قیمت" موجود در دیتاست اولیه، به منظور تعیین ویژگی مجموع مبالغ سفارش دهی مشتری در هر بار مراجعه استفاده شده است. هدف از ایجاد این

- 
1. Customer Frequency 1397
  2. Total Customer Order Num Per Visit
  3. Total Customer Order Num Per Visit
  4. Total Customer Order Num Per Visit
  5. Burger
  6. Sandwich
  7. Fried
  8. Pizza
  9. Appetizer
  10. Desert
  11. Salad
  12. Natural Drink
  13. Carbonated Drink

ویژگی، ایجاد ۳ ویژگی مجموع مبالغ سفارش دهی هر مشتری در سال‌های ۱۳۹۵، ۱۳۹۶ و ۱۳۹۷ بوده است که از طریق فرمول نویسی در اکسل ایجاد شده‌اند. به‌طور کلی، در انتهای مرحله خلق ویژگی، تعداد ۳۵ ویژگی ایجاد شدند که در جدول شماره ۲ ارائه شده‌اند:

جدول ۲. فهرستی از فیلدهای اطلاعاتی دیتاست

ردیف	ویژگی	نوع داده	توضیحات
۱	Customer ID	Int	شناسه مشتری
۲	Customer Name	String	نام مشتری
۳	Address	String	آدرس
۴	Region Name	String	نام منطقه
۵	Region ID	Int	شناسه منطقه
۶	Order Date	Date	تاریخ سفارش دهی
۷	Order Date Year	Date	تاریخ سفارش دهی (سال)
۸	Order Date Month	Date	تاریخ سفارش دهی (ماه)
۹	Order Date Day	Date	تاریخ سفارش دهی (روز)
۱۰	Branch Name	String	نام شعبه
۱۱	Branch Code	Int	کد شعبه
۱۲	Product Code	Int	کد محصول
۱۳	Product Name	String	نام محصول
۱۴	Product Category	String	گروه محصول
۱۵	Burger	Binary	آیا مشتری گروه "برگر" را انتخاب کرده است؟
۱۶	Pizza	Binary	آیا مشتری گروه "پیتزا" را انتخاب کرده است؟
۱۷	Appetizer	Binary	آیا مشتری گروه "پیش‌غذا" را انتخاب کرده است؟
۱۸	Desert	Binary	آیا مشتری گروه "دسر" را انتخاب کرده است؟

ردیف	ویژگی	نوع داده	توضیحات
۱۹	Salad	Binary	آیا مشتری گروه "سالاد" را انتخاب کرده است؟
۲۰	Sandwich	Binary	آیا مشتری گروه "ساندویچ" را انتخاب کرده است؟
۲۱	Fried	Binary	آیا مشتری گروه "سوخاری" را انتخاب کرده است؟
۲۲	Carbonated Drink	Binary	آیا مشتری گروه "نوشیدنی گازدار" را انتخاب کرده است؟
۲۳	Natural Drink	Binary	آیا مشتری گروه "نوشیدنی طبیعی" را انتخاب کرده است؟
۲۴	Customer Order Num Per Visit	Int	تعداد سفارش‌های مشتری در هر بار مراجعه
۲۵	Customer Frequency 1395	Int	تعداد دفعات مراجعه مشتری در سال ۱۳۹۵
۲۶	Customer Frequency 1396	Int	تعداد دفعات مراجعه مشتری در سال ۱۳۹۶
۲۷	Customer Frequency 1397	Int	تعداد دفعات مراجعه مشتری در سال ۱۳۹۷
۲۸	Total Customer Order Num 1395	Int	تعداد کل سفارش‌های هر مشتری در سال ۱۳۹۵
۲۹	Total Customer Order Num 1396	Int	تعداد کل سفارش‌های هر مشتری در سال ۱۳۹۶
۳۰	Total Customer Order Num 1397	Int	تعداد کل سفارش‌های هر مشتری در سال ۱۳۹۷
۳۱	Production Price	Int	قیمت هر محصول
۳۲	Total Customer Order Price Per Visit	Int	کل مبلغ سفارش دهی هر مشتری در هر بار مراجعه <sup>۱</sup>

۱. این ویژگی از حاصل ضرب "تعداد" در "قیمت واحد محصول" به دست آمده است.

ردیف	ویژگی	نوع داده	توضیحات
۳۳	Total Order Amount Per Customer 1395	Int	کل مبلغ سفارش دهی هر مشتری در سال ۱۳۹۵
۳۴	Total Order Amount Per Customer 1396	Int	کل مبلغ سفارش دهی هر مشتری در سال ۱۳۹۶
۳۵	Total Order Amount Per Customer 1397	Int	کل مبلغ سفارش دهی هر مشتری در سال ۱۳۹۷

همان‌طور که پیش‌تر ذکر گردید، به دلیل ایجاد قابلیت رصد مشتریان، فقط اطلاعات ذخیره‌شده از مشتریانی که اشتراک دریافت کرده و "تحويل در محل" بوده‌اند، مورد بررسی قرار گرفته است. بدین ترتیب، هر مشتری فقط با یک کد شناسایی شده است.

در مرحله دوم، ویژگی‌های تکراری و فاقد اهمیت مانند "نام مشتری"، "آدرس"، "نام منطقه"، "تاریخ سفارش دهی"، "نام شعبه"، "نام محصول"، "گروه محصول"، "تعداد سفارش‌های مشتری در هر بار مراجعه"، "قیمت هر محصول" و "کل مبلغ سفارش دهی هر مشتری در هر بار مراجعه" حذف شدند و تعداد ۲۵ فیلد اطلاعاتی نهایی مورد بررسی قرار گرفتند.

در مرحله سوم، از طریق کد نویسی به زبان پایتون رکوردهای تکراری شناسایی و حذف شدند. علاوه بر حذف رکوردهای تکراری، دادگان پرت نیز به روش Z-Score شناسایی و در خوشه جداگانه‌ای قرار داده شدند. تعداد ۵۱۶۵۰ رکورد در خوشه دادگان پرت قرار گرفتند. لازم به ذکر است که هیچ مورد مقادیر ازدست‌رفته در این مجموعه داده وجود نداشته است.

در طول این پروسه نیاز به نرمال‌سازی داده‌ها برای تطبیق‌پذیری بهتر با خوشه بندها احساس می‌شد که از طریق کد نویسی به زبان پایتون بر روی داده‌ها انجام شده است. همچنین در راستای کاهش فضای ذخیره‌سازی و حجم پردازش داده‌ها، الگوریتم PCA از طریق کد نویسی به زبان پایتون بر روی داده‌ها اعمال گردیده است و نتایج مجدداً

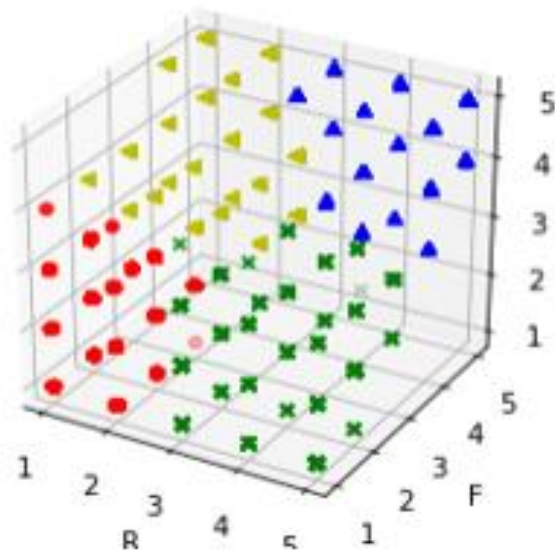
ذخیره‌سازی شده‌اند. سپس رکوردها تجمیع<sup>۱</sup> شده‌اند به طوری که هر رکورد از دیتاست، دقیقاً مربوط به یک کاربر بوده است.

پس از اعمال عملیات پاک‌سازی داده، یک و نیم میلیون رکورد داده به تعداد ۳۲۷۷۸۲ رکورد داده کاهش داده شده است که اطلاعات مشتریان رستوران زنجیره‌ای را ارائه کرده است.

### مدل‌سازی (خوشه‌بندی مشتریان)

در این تحقیق، چهار روش خوشه‌بندی K-Means، K-Means+، Mean Shift و SOM برای شناسایی گروه‌های مشتریان مورد استفاده قرار گرفته است. در این روش‌های خوشه‌بندی از داده‌های موجود در فیله‌های "تعداد دفعات مراجعه مشتری" و "مجموع مبالغ سفارش دهی" در ۳ سال ۱۳۹۵، ۱۳۹۶ و ۱۳۹۷ که کمی و پیوسته هستند استفاده شده و فیلد "آخرین تاریخ مراجعه" نیز به فیلد کمی پیوسته تبدیل شده و مورد استفاده قرار گرفته است. در جریان مدل‌سازی خوشه‌بندی از روش RFM استفاده شده است. در پیاده‌سازی روش RFM به زبان پایتون، این سه فاکتور به اعداد یک تا پنج نگاشت شدند و سپس بر اساس نتایج به دست آمده، خوشه‌بندی انجام شده است. در شکل شماره "۲" نمونه‌ای از نمودار نقطه‌ای پراکنندگی سه معیار "M و F، R" در الگوریتم K-Means+ ترسیم شده است:



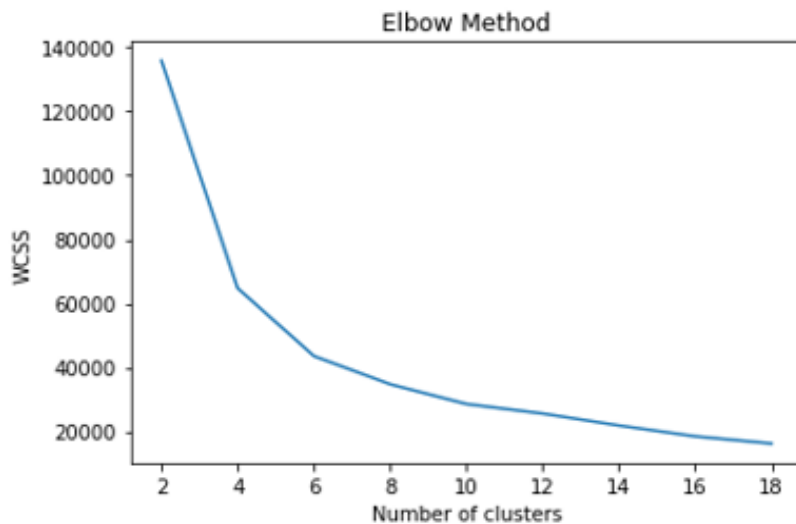


شکل ۲. نمودار نقطه‌ای پراکنده‌ی سه معیار "F, R" و "M" در الگوریتم K-Means

در این تحقیق، به منظور تعیین تعداد خوشه‌های بهینه از روش Elbow<sup>۱</sup> استفاده شده است. معیار ارزیابی میزان خطا در روش Elbow مجذور فاصله اقلیدسی<sup>۲</sup> بوده است. تعداد خوشه بهینه استخراجی از الگوریتم K-Means با استفاده از روش Elbow برابر با ۸ خوشه بوده است. شکل شماره "۳" نمودار میزان خطا و تعداد خوشه بهینه پس از به کارگیری روش Elbow را نمایش داده است:

---

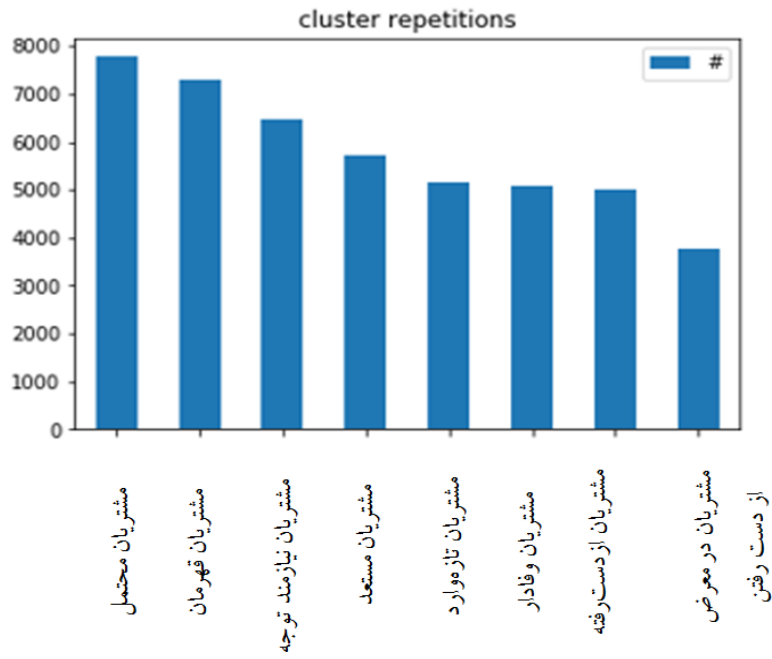
1. Elbow Method  
2. WCSS



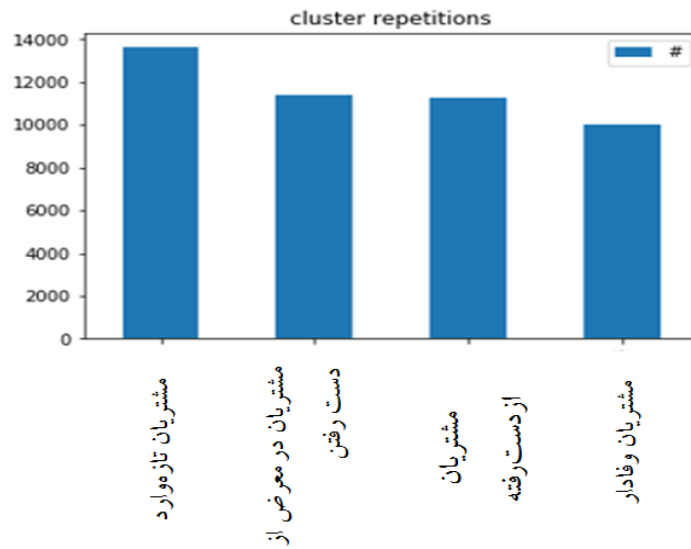
شکل ۳. نمودار میزان خطا و تعداد خوشه بهینه بر اساس روش **Elbow**

### نتایج خوشه‌بندی

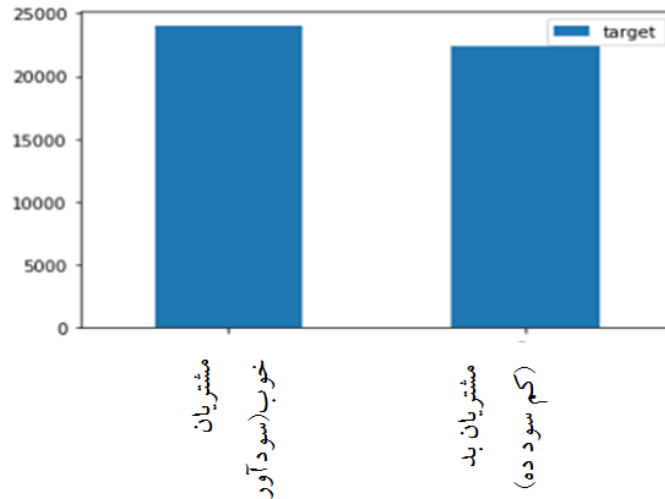
پس از پیاده‌سازی الگوریتم‌های خوشه‌بندی بر روی داده‌های پاک‌سازی شده، بر اساس الگوریتم K-Means، مشتریان رستوران زنجیره‌ای در ۸ خوشه، بر اساس الگوریتم K-Means+ در ۴ خوشه، بر اساس الگوریتم SOM در ۴ خوشه و بر اساس الگوریتم Mean-Shift مشتریان رستوران زنجیره‌ای در ۲ خوشه، قرار داده شدند. نتایج به‌دست آمده از پیاده‌سازی الگوریتم‌های خوشه‌بندی به همراه عنوان در نظر گرفته شده برای هر خوشه و فراوانی خوشه‌ها، در نمودارهای ذیل ترسیم شده است:



شکل ۴. نتایج حاصل از پیاده‌سازی الگوریتم K-Means



شکل ۵. نتایج حاصل از پیاده‌سازی الگوریتم K-Means+



شکل ۷. نتایج حاصل از پیاده‌سازی الگوریتم Mean-Shift

به دلیل آنکه تعداد خوشه‌های ایجاد شده حاصل از پیاده‌سازی الگوریتم K-Means با تعداد خوشه‌های بهینه تعیین شده از طریق روش Elbow برابر بوده است و کمترین خطای خوشه‌بندی را داشته است، در این تحقیق به‌عنوان الگوریتم برتر خوشه‌بندی انتخاب و به‌عنوان مرجع در نظر گرفته شده است. در اصل، به‌کارگیری روش Elbow به‌منظور تعیین خطای خوشه‌بندی به‌عنوان نوآوری این تحقیق در نظر گرفته شده است.

#### تفسیر نتایج خوشه‌بندی

پس از تعیین تعداد خوشه‌های مشتریان رستوران زنجیره‌ای، الگوهای رفتاری مشتریان این خوشه‌ها بر اساس سه معیار "تازگی خرید"<sup>۱</sup>، "تعداد دفعات خرید"<sup>۲</sup> و "مبلغ خرید"<sup>۳</sup> شرح داده شده است و استراتژی‌های بازاریابی، جذب و ابقاء مشتریان با توجه به رفتار هر خوشه تدوین شده است که در جدول شماره "۳" ارائه شده است:

1. Recency
2. Frequency
3. Monetary

جدول ۳. الگوی رفتاری خوشه مشتریان بر اساس الگوریتم K-Mean و استراتژی‌های بازاریابی،

جذب و ابقاء مشتریان

خوشه	استراتژی
مشتریان محتمل	بر اساس نتایج به‌دست‌آمده از روش RFM، این خوشه از مشتریان، اخیراً به رستوران مراجعه کرده‌اند اما مجموع مبالغ سفارش دهی آن‌ها کم بوده است. این احتمال وجود دارد که از طریق برقراری روابط با اعضای این خوشه، درصدی از این اعضا به مشتریان وفادار تبدیل شوند. از این رو، پیشنهاد شده است که به‌منظور جذب هر چه بیشتر این اعضا، از طریق ارسال پیامک‌های تبلیغاتی به آن‌ها، مقدمات آشنایی هر چه بیشتر این اعضا با رستوران فراهم شود. همچنین، از طریق ارائه رایگان محصولات این رستوران به اعضا این خوشه، با هدف آشنایی بیشتر آن‌ها پروسه جذب آن‌ها بهبود داده می‌شود.
مشتریان نیازمند توجه	بر اساس نتایج به‌دست‌آمده از روش RFM، مشتریان موجود در این خوشه مدت‌زمان زیادی به رستوران مراجعه نکرده‌اند اما بر اساس سوابق ثبت شده از این مشتریان، تعداد دفعات مراجعه آن‌ها نسبتاً زیاد بوده است و مبالغ سفارش دهی نسبتاً خوبی داشته‌اند؛ بنابراین، پیشنهاد شده است که به‌منظور جذب بیشتر این اعضا، در بازه‌های زمانی کوتاه‌مدت، تخفیف‌هایی به این اعضا ارائه شود. همچنین، محصولاتی که در سوابق سفارش دهی آن‌ها بیشتر دیده شده است و جزء علاقه‌مندی‌های این اعضا بوده است، به آن‌ها پیشنهاد شود، بهتر است این محصولات با در نظر گرفتن تخفیف ارائه گردد.
مشتریان از دست‌رفته	این خوشه از مشتریان، پایین‌ترین مبلغ خرید را در بین مشتریان داشته‌اند و از آخرین تاریخ مراجعه آن‌ها زمان زیادی گذشته است. این خوشه از مشتریان، عملاً رستوران زنجیره‌ای را حذف کرده‌اند؛ بنابراین پیشنهاد شده است که از طریق برگزاری کمپین، مجدداً این مشتریان به رستوران بازگردانده شوند و در صورتی که این روش مفید واقع نشود، از این مشتریان صرف‌نظر گردد.
مشتریان در معرض از دست رفتن	تعداد دفعات مراجعه، تازگی و مبلغ سفارش دهی مشتریان این خوشه کمتر از میانگین نتایج ثبت شده از سایر مشتریان بوده است؛ بنابراین پیشنهاد شده است که به‌منظور جلوگیری از خروج این دسته از مشتریان از لیست مشتریان رستوران، محصولاتی که ارزش خرید بیشتری از دیدگاه مشتریان دارند به آن‌ها پیشنهاد شود، محصولات جدید رستوران در نظر گرفتن تخفیف به آن‌ها ارائه گردد، ایمیل‌ها/ پیامک‌های شخصی به‌منظور برقراری ارتباط مجدد با این اعضا و همچنین تکریم آن‌ها ارسال شود.
مشتریان قهرمان	مشتریان موجود در این خوشه اغلب به رستوران مراجعه کرده‌اند، از آخرین تاریخ مراجعه آن‌ها مدت‌زمان زیادی نمی‌گذرد و بالاترین مبالغ سفارش دهی را داشته‌اند. این خوشه از مشتریان بالاترین درصد سودآوری رستوران زنجیره‌ای را به خود اختصاص داده‌اند؛

خوشه	استراتژی
	بنابراین، پیشنهاد می‌شود که محصولات با قیمت بالاتر به این مشتریان پیشنهاد داده شود و از این طریق سودآوری رستوران افزایش داده شود.
مشتریان مستعد	مشتریان موجود در این خوشه، اخیراً به رستوران مراجعه کرده‌اند و مبالغ سفارش دهی نسبتاً خوبی پرداخت کرده‌اند. پیشنهاد شده است به منظور تبدیل این اعضاء به مشتریان وفادار، پیشنهاد عضویت در باشگاه مشتریان و برخورداری از مزایای این باشگاه (به‌عنوان مثال تخفیف‌هایی با درصد بالا) ارائه شود و محصولات جدید رستوران به این اعضاء معرفی و ارائه گردد.
مشتریان تازه‌وارد	این خوشه از مشتریان، به تازگی به رستوران مراجعه کرده بودند اما تعداد دفعات مراجعه آن‌ها کم بوده است. پیشنهاد شده است که به منظور جذب هر چه بیشتر این خوشه از مشتریان، در بازه‌های زمانی برنامه‌ریزی شده، تخفیف‌هایی به این اعضاء ارائه شود، از آخرین تجربیات خرید این اعضاء بازخورد گرفته شود و از این طریق روابطی با این اعضاء پایه‌گذاری شود.
مشتریان وفادار	این خوشه از مشتریان، همراهان همیشگی رستوران بوده‌اند و از ویژگی‌های بارز اعضای این خوشه این بوده است که در قبال نظرسنجی‌ها و بهبود عملکرد رستوران پاسخگو بوده‌اند؛ بنابراین، پیشنهاد شده است از طریق مشارکت دادن این اعضا در نظرسنجی عملکرد رستوران، شعب رستوران و یا محصولات رستوران، رابطه قدرتمندی با این اعضاء برقرار گردد و از طریق تکریم آن‌ها موجبات جلب رضایت بیشتر این اعضاء فراهم شود.

### مدل‌سازی (دسته‌بندی مشتریان)

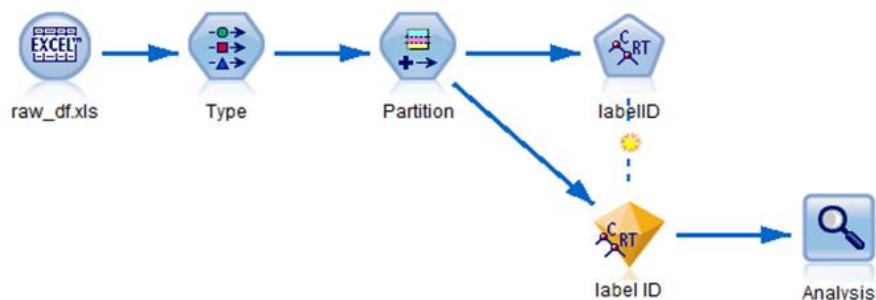
در دیتاست ایجاد شده پس از پیاده‌سازی الگوریتم K-Means ویژگی تحت عنوان "نام خوشه" ایجاد شده است و به وضوح مشخص شده است که کدام مشتری در چه خوشه‌ای قرار داده شده است. در این تحقیق، هدف از مدل‌سازی دسته‌بندی، کشف الگوی رفتاری مشتریان خوب (سودآور) و بد (کم سود ده) و ایجاد مدلی به منظور پیش‌بینی رفتار آتی مشتریان جدید بوده است.

ویژگی Label ID به‌عنوان فیلد هدف در نظر گرفته شده است. این فیلد داده‌ای از نوع باینری بوده و مقدار ۱ نشان‌دهنده مشتری خوب (سودآور) بوده است. در این مرحله،

پس از کسب نظر از خبره کسب و کار، ویژگی‌های مشتری خوب (سودآور) استخراج شده است و با نتایج حاصل از خوشه‌بندی ارتباط داده شده است.

مدل‌سازی دسته‌بندی با استفاده از نرم‌افزار IBM Spss Modeler<sup>۱</sup> بر اساس چهار الگوریتم Chaid، C5، C&R و Quest انجام شده است. ۳۰ درصد از داده‌ها به‌عنوان داده آموزشی و ۷۰ درصد از داده‌ها به‌عنوان داده آزمایشی در نظر گرفته شده است.

لازم به ذکر است، یکی از اهداف این پژوهش مقایسه الگوریتم‌های دسته‌بندی مختلف از نظر میزان صحت و دقت در نتایج و تعداد قواعد استخراجی است که در ادامه در جدول ۵ قابل مشاهده است. به همین دلیل دسته‌بندی خودکار<sup>۲</sup> انجام نشده است. در ذیل نمونه‌ای از مدل‌سازی دسته‌بندی در شکل "۸" ترسیم شده است:



شکل ۸. نمونه‌ای از مدل طبقه‌بندی در نرم‌افزار IBM Spss Modeler

فیلدهای اطلاعاتی موجود در دیتاست در چهار گروه به‌صورت دوبه‌دو باهم در نظر گرفته شده‌اند و هر بار یک الگوریتم بر روی داده‌ها اعمال شده است. در جدول شماره "۴" گروه‌بندی فیلدهای اطلاعاتی ارائه شده است:

- 
1. Clementine
  2. Auto classifier

جدول ۴. گروه‌بندی فیلهای اطلاعاتی

فیلهای آدرس				
(۱)	کد شعبه	(۲)	کد منطقه	
فیلهای نوع سفارش				
(۳)	پیتزا؟	(۴)	پیش‌غذا؟	(۵)
(۶)	سوخاری؟	(۷)	ساندویچ؟	(۸)
(۹)	نوشیدنی کارخانه‌ای؟	(۱۰)	دسر؟	(۱۱)
فیلهای تعداد دفعات مراجعه				
(۱۲)	تعداد دفعات مراجعه در سال ۱۳۹۵	(۱۳)	تعداد دفعات مراجعه در سال ۱۳۹۶	(۱۴)
فیلهای تعداد سفارش‌ها				
(۱۵)	تعداد سفارش‌ها در سال ۱۳۹۵	(۱۶)	تعداد سفارش‌ها در سال ۱۳۹۶	(۱۷)
فیلهای مجموع مبلغ سفارش دهی				
(۱۸)	مجموع مبلغ سفارش دهی در سال ۱۳۹۵	(۱۹)	مجموع مبلغ سفارش دهی در سال ۱۳۹۶	(۲۰)

بدین ترتیب، ۲۸ مدل ایجاد شده است که تعدادی قواعد رفتاری از هر یک از این مدل‌ها استخراج شده است. لازم به ذکر است تعداد قواعد استخراج شده از هر مدل و همچنین دقت و صحت هر مدل در جدول شماره "۵" ارائه شده است:

جدول ۵. نتایج دسته‌بندی مشتریان

مدل	فیلهای آدرس	فیلهای نوع سفارش	فیلهای تعداد دفعات مراجعه	فیلهای تعداد سفارش‌ها	فیلهای مبلغ سفارش دهی	فیلد هدف	تعداد قواعد استخراج شده	صحت و دقت
Chaid	*				*	Lable ID	۲۰	۹۵/۲۶
Quest	*				*	Lable ID	۵	۹۵/۸۵
C5	*				*	Lable ID	۱۸	۹۶/۶
C&R	*				*	Lable	۶	۹۶/۵



صحت و دقت	تعداد قواعد استخراج شده	فیلد هدف	فیلدهای مبلغ سفارش دهی	فیلدهای تعداد سفارش‌ها	فیلدهای تعداد دفعات مراجعه	فیلدهای نوع سفارش	فیلدهای آدرس	مدل
		ID						
۹۱/۶۸	۸	Lable ID				*	*	Chaid
۹۱/۶۸	-	Lable ID				*	*	Quest
۹۱/۶۸	-	Lable ID				*	*	C5
۹۱/۶۸	-	Lable ID				*	*	C&R
۹۶/۰۲	۱۲	Lable ID			*		*	Chaid
۹۷/۸۳	۷	Lable ID			*		*	Quest
۹۸/۶۱	۱۱	Lable ID			*		*	C5
۹۸/۴۲	۶	Lable ID			*		*	C&R
۹۵/۲۶	۲۰	Lable ID		*			*	Chaid
۹۶/۰۴	۶	Lable ID		*			*	Quest
۹۷/۴۵	۷	Lable ID		*			*	C5
۹۷/۲۷	۶	Lable ID		*			*	C&R
۹۵/۹۳	۱۳	Lable ID		*	*			Chaid
۹۶/۸۳	۵	Lable ID		*	*			Quest
۹۸/۱۴	۶	Lable ID		*	*			C5
۹۸/۰۳	۵	Lable ID		*	*			C&R
۹۵/۳۵	۱۴	Lable ID		*		*		Chaid
۹۶/۶۳	۶	Lable ID		*		*		Quest
۹۷/۲۳	۷	Lable ID		*		*		C5

مدل	فیلدهای آدرس	فیلدهای نوع سفارش	فیلدهای تعداد دفعات مراجعه	فیلدهای تعداد سفارشات	فیلدهای مبلغ سفارش دهی	فیلد هدف	تعداد قواعد استخراج شده	صحت و دقت
C&R		*		*		Lable ID	۶	۹۷/۱۱
Chaid		*	*			Lable ID	۱۳	۹۶/۰۲
Quest		*	*			Lable ID	۵	۹۶/۸۳
C5		*	*			Lable ID	۱۳	۹۸/۲۴
C&R		*	*			Lable ID	۵	۹۸/۰۳

### قواعد استخراج شده

قواعد استخراج شده از مدل‌سازی‌های صورت گرفته، به زبان ساده و با هدف درک آسان توسط صاحبان کسب و کار تفسیر شده‌اند. این الگوها عبارت‌اند از:

جدول ۶. قواعد استخراج شده از الگوریتم‌های دسته‌بندی

شماره قاعده	دسته مشتری	کل مبلغ خرید مشتری (۱۳۹۷)	کل مبلغ خرید مشتری (۱۳۹۶)	کل مبلغ خرید مشتری (۱۳۹۵)	محل سکونت (منطقه) (۱)	محل سکونت (منطقه) (۳)	محل سکونت (منطقه) (۴)
۱	خوب	۹۶۹۲۵۰ <					
۲	خوب	۹۶۹۲۵۰ >	۴۱۶۸۷۵ <				
۳	خوب	۹۶۹۲۵۰ >		۳۲۰۶۸۷ <			
۴	خوب	۶۰۶۶۲۶ < ۹۶۹۲۵۰ >			✓		
۵	خوب		۵۸۵۵۰۰ <				
۶	خوب		۵۸۵۵۰۰ >	۴۵۳۰۰۰ <			
۷	بد			۳۲۰۶۸۷ >			
۸	بد	۶۰۶۶۲۶ < ۹۶۹۲۵۰ >				✓	✓

جدول ۷. قواعد استخراج شده از الگوریتم‌های دسته‌بندی

شماره قاعده	دسته مشتری	تعداد دفعات مراجعه مشتری (۱۳۹۷)	تعداد دفعات مراجعه مشتری (۱۳۹۶)	تعداد دفعات مراجعه مشتری (۱۳۹۵)	محل سکونت (منطقه ۱)	محل سکونت (منطقه ۳)	محل سکونت (منطقه ۴)
۹	خوب	۵۰<					
۱۰	خوب	۵۰>	۳۰<				
۱۱	خوب	۴۰>	۳۰<	۳۰<			
۱۲	خوب		۳۰<	۳۰<	✓		
۱۳	خوب	۴۰>	۳۰<	۳۰<		✓	✓
۱۴	بد		۳۰>			✓	✓

جدول ۸. قواعد استخراج شده از الگوریتم‌های دسته‌بندی

شماره قاعده	دسته مشتری	تعداد کل سفارش‌ها (۱۳۹۷)	تعداد کل سفارش‌ها (۱۳۹۶)	تعداد کل سفارش‌ها (۱۳۹۵)	محل سکونت (منطقه ۱)	محل سکونت (منطقه ۳)	محل سکونت (منطقه ۴)
۱۵	خوب	۵۵۰<					
۱۶	خوب	۵۵۰>	۲۵۰<				
۱۷	خوب	۵۵۰>		۲۵۰<			
۱۸	خوب	۳۵۰<	۲۵۰>	۲۵۰>	✓		
۱۹	بد	۳۵۰>	۲۵۰>	۲۵۰>			
۲۰	بد	۳۵۰<	۲۵۰>	۲۵۰>		✓	✓

### رتبه‌بندی قواعد استخراج شده

در گام آخر، به منظور اثبات روایی و پایایی مدل دسته‌بندی، قواعد استخراج شده حاصل از اجرای مدل‌ها در اختیار گروهی متشکل از ۸ نفر از مدیران رستوران زنجیره‌ای رستوران قرار گرفتند. این گروه متشکل از مدیرعامل مجموعه رستوران، مدیر هریک از شعب، مدیر فناوری ارتباطات و مدیر بخش CRM یا مدیریت ارتباط با مشتریان بوده است. ابتدا، در خصوص نحوه ایجاد قواعد و مدل‌سازی توضیحاتی ارائه گردید و سپس معیارهای

ارزیابی کارایی این قواعد توسط مدیران تعیین شده‌اند. چهار معیار "میزان کاربردی بودن"، "میزان اهمیت شعب"، "میزان اهمیت برای واحد CRM" و "سهولت بررسی" به عنوان معیارهای ارزیابی از سوی مدیران انتخاب شدند. در مرحله بعد، هر یک از مدیران به قواعد استخراج شده بر اساس معیارهای ارزیابی کارایی که پیش‌تر تعیین و مورد تأیید قرار گرفته بود، امتیازدهی کردند.

تمامی امتیازها جمع‌آوری شدند و رتبه‌بندی قواعد با استفاده از روش SAW که یکی از روش‌های MADM است، انجام شد. نتایج رتبه‌بندی قواعد مطابق در جدول "۶" ارائه شده است:

جدول ۹. رتبه‌بندی قواعد استخراج شده

رتبه	قاعده	رتبه	قاعده
۱	قاعده (۱۱)	۲	قاعده (۱۲)
۳	قاعده (۱۶)	۴	قاعده (۱۴)
۵	قاعده (۱۵)	۶	قاعده (۱۰)
۷	قاعده (۱۳)	۸	قاعده (۸)
۹	قاعده (۹)	۱۰	قاعده (۱۷)
۱۱	قاعده (۱)	۱۲	قاعده (۲)
۱۳	قاعده (۳)	۱۴	قاعده (۴)
۱۵	قاعده (۵)	۱۶	قاعده (۶)
۱۷	قاعده (۷)	۱۸	قاعده (۸)
۱۹	قاعده (۱۹)	۲۰	قاعده (۲۰)

پس از استخراج قواعد و شناسایی مشتریان در دسته‌های استخراج شده، این اطلاعات در اختیار امور مشتریان و CRM رستوران قرار گرفت و با تعبیه مکانیسم‌های مناسب بازاریابی برای مشتریان خوب (وفادار) به‌ویژه در دوران کرونا، با وجود افت بسیار زیاد فروش در این دوران، این دسته را ترغیب نمودند که از محصولات سالم و با رعایت پروتکل‌های بهداشتی رستوران بتوانند با اطمینان کافی سفارش دهند. طبیعی است انرژی گذاشتن فقط

برای این دسته از مشتریان، در هدفمند نمودن تبلیغات و بازاریابی بسیار مؤثر و هزینه‌های بازاریابی را حداقل به میزان ۳۰ درصد کاهش داده است.

### محدودیت و نوآوری

در این تحقیق از روش Elbow به منظور مشخص کردن خطای خوشه‌بندی استفاده شده است که در تحقیقاتی از این دست، این روش کمتر مورد استفاده قرار گرفته است.

از جمله نوآوری‌های این تحقیق، می‌توان به به‌کارگیری تکنیک SAW که یکی از تکنیک‌های کاربردی در روش تصمیم‌گیری چند معیاره است، اشاره نمود. قواعد استخراج شده حاصل از پیاده‌سازی الگوریتم‌های دسته‌بندی بر روی داده‌های مشتریان با استفاده از این تکنیک مورد ارزیابی قرار گرفتند و رتبه‌بندی شدند.

همچنین، در این تحقیق از نتایج به‌دست‌آمده از فاز مدل‌سازی خوشه‌بندی به منظور دسته‌بندی مشتریان در فاز دسته‌بندی نیز استفاده شده است و بدین ترتیب بین دو فاز مدل‌سازی ارتباط برقرار شده است.

از مهم‌ترین محدودیت‌های این پژوهش، نبود شناسه منحصر به فرد به منظور شناسایی مشتریان و ره‌گیری عملکرد و سوابق آن‌ها بوده است. از این رو، به رستوران زنجیره‌ای پیشنهاد شده است که از سیستم شناسایی منحصر به فرد به منظور شناسایی مشتریان خود در شعب مختلف، استفاده کنند. همچنین، در این پژوهش، فقط بر روی داده مشتریانی که سفارش خود را به صورت تلفنی یا از طریق نرم‌افزارهای فروش آنلاین غذا سفارش داده‌اند، تحلیل صورت گرفته است و سوابق کامل مشتری در تعداد دفعات حضور در سالن در دسترس نبوده است. از دیگر محدودیت‌های این تحقیق، حجم بالای داده‌ها و در دسترس نبودن تجهیزات و سیستم‌های پیشرفته به منظور تحلیل و بررسی داده‌ها بوده است.

### نتیجه‌گیری و پیشنهادها

الگوریتم خوشه‌بندی K-Means و روش RFM نتایج بسیار مفید و تأثیرگذار را در شرکت‌های خدماتی مانند رستوران‌ها، هتل‌ها و غیره قرار می‌دهند. به‌کارگیری روش‌های

داده‌کاوی در این نوع شرکت‌ها، ضمن کمک به اهداف بازاریابی و تبلیغات، ارتباط مجموعه با مشتریان را نیز بهبود می‌بخشد و به افزایش سودآوری کمک می‌کند. با توجه به اینکه این تحقیق با استفاده از روش RFM انجام شده است و از اطلاعات دموگرافیک مشتریان استفاده نشده است، پیشنهاد می‌شود که اطلاعات دموگرافیک مشتریان رستوران ثبت شود تا در تحقیقات آتی مورد استفاده قرار گیرد. همچنین، نظرسنجی‌های مشتریان و امتیازات ارائه شده به عملکرد رستوران جمع‌آوری شود و از این نظرسنجی‌ها نیز به عنوان یک شاخص ارزیابی به صورت ترکیبی با سایر مشخصه‌ها به منظور بررسی عملکرد شعب مختلف رستوران استفاده شود.

#### ORCID

Fatemeh Rahimi



<http://orcid.org/>

Mohammad Vahid Sebt



<http://orcid.org/>

Nasim Ghanbar Tehrani



<http://orcid.org/>

## منابع

- آهنگری، سعیده سادات؛ مؤمنی، منصور و جعفر نژاد، احمد. (۱۳۹۲). بخش‌بندی مشتریان بانک شهر به وسیله شبکه‌های عصبی SOM. اولین کنفرانس بین‌المللی تکنیک‌های مدیریت و حسابداری، تهران.
- صدوقی نیا، الهام؛ الهی، شعبان؛ حسن زاده، علیرضا و صحرائی، شقایق. (۱۳۹۸). سیستم خیره ارزشیابی وفاداری مشتریان در صنعت پخش و توزیع. نشریه علمی مطالعات مدیریت کسب‌وکار هوشمند، ۸(۳۰)، ۷۸-۵۵.
- فروزنده، سمیه. (۱۳۹۶). پیش‌بینی رویگردانی مشتریان حقیقی بانک ملت با استفاده از روش‌های جمعی مبتنی بر بگینگ و بوستینگ. پایان‌نامه کارشناسی ارشد، دانشگاه خوارزمی.
- قاسم‌زاده، سیده نفیسه. (۱۳۹۵). بخش‌بندی مشتریان اینترنت شبکه تلفن همراه و بررسی تغییرات رفتاری آن‌ها. پایان‌نامه کارشناسی ارشد، دانشگاه خوارزمی.
- کدخدایی، الهه. (۱۳۹۴). کشف دانش از الگوهای متوالی RFM برای بخش‌بندی مشتریان. پایان‌نامه کارشناسی ارشد، دانشگاه خوارزمی.
- محمدی، شبنم و علیزاده، سمیه. (۱۳۹۳). تحلیل مشکلات شعب بانک آینده در سراسر کشور با استفاده از روش داده‌کاوی. مدیریت فناوری اطلاعات، ۶(۲)، ۳۵۰-۳۳۳.
- موتمنی، علیرضا؛ مرادی، هادی و همتی، امین. (۱۳۹۱). شناسایی و رتبه‌بندی عوامل مؤثر بر وفاداری خرده‌فروشان به شرکت‌های توزیع. فصل‌نامه علمی ترویجی مدیریت زنجیره تأمین، ۱۴(۳۵)، ۷۵-۶۲.

## References

- Bach, M.P., Jankovic, S., Dominic, K., & Sarlija, N. (2013). business client segmentation using self - organizing map. *South East European journal of Economics and Business*, 8(2), 32-41.
- Berry, M., & Linoff, G. (2004). *Data Mining Techniques For Marketing, Sales and Customer Relationship Management Second Edition*. Wiley Publishing, Inc., Indianapolis, Indiana.
- Burnett, A. (2001). *Handbook of Key Customer Relationship Management: The Definitive Guide to Winning, Managing and Developing Key Account business*. Pearson Education.

- Burnett, A. (2001). *Handbook of Key Customer Relationship Management: The Definitive Guide to Winning, Managing and Developing Key Account business* Pearson Education.
- Chadors, A. J. & V.E.D. (2014). *customer segmentation model based on value generation for marketing strategy foundation*. Estudios Gerenciales, 30, 25-30.
- Cheng, C. H., & Chen, Y. S. (2009). classifying the segmentation of customer value via RFM theory. *Expert systems with applications*,36(3), 4176-4184.
- Dursun & Caber, M (2016) Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives*, 18, 153-160
- Hand, D., Mannila, H., & Smyth, P. (2001). *principles of Data mining*. The MIT press.
- Huang, S-C., Chang, E-C. & Wu, H-H. (2009). A case study of applying data mining techniques in an outfitters customer value analysis. *Expert Systems with Applications*, 36, 5909-5915.
- Liang, Y-H. (2010). Integration of data mining technologies to analyze customer value for the automotive maintenance industry. *Expert Systems with Applications*, 37, 7489-7496.
- Liao, S-H., Chen, Y-N. & Tseng, Y-Y. (2009). mining demand chain knowledge of life insurance market for new product development. *Expert systems with application*, 36(5), 9422-9437
- Ngai, E. W. T., Xiu, L., & Chau, D.C.K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *expert systems with applications*, 36, 2592-2602.
- Nioosha jafari momtaz, somaye alizade, mahyar sharifi vaghefi, (2013). A new model for assessment fast food customer behavior, *emerald insight*, 313, 601-613.
- Olson, D. L., & Chae, B. K. (2012). Direct marketing decision support through predictive customer response modeling. *Decision Support Systems*, 54(1), 443-451.
- Rygielski, C., Wang, J-C. & Yen, D. (2002). Data mining technologies for customer relationship management. *Technology in society*, 24, 438-502.

### References [In Persian]

- Ahangari, S. S., Momeni, M., & Jafarnejad, A. (2013). *Bank Shahr customers' segmentation by SOM neural networks*. The First International Conference on Management and Accounting Techniques, Tehran.[In Persian]



- Sadoughinia, E., Hasanzadeh, A., Elahi, Sh., & Sahraei, Sh. (2020). An Expert System for Evaluating Customers' Loyalty in Distribution Industry. *Quartly Journal of BI Management Studies*, 8(30), 55-78. [In Persian]
- Forouzandeh, S. (1396). *Predicting the turnover of Bank Mellat's real customers using collective methods based on bagging and boosting*. (Master Thesis), Kharazmi University. [In Persian]
- Qasemzadeh, S. N. (1395). *Segmenting Internet network mobile customers and examining their behavioral changes*. (Master Thesis), Kharazmi University. [In Persian]
- Kadkhodai, E. (1394). *Discover knowledge of successive RFM patterns for customer segmentation*. (Master Thesis), Kharazmi University. [In Persian]
- Mohammadi, Sh., and Alizadeh, S. (2014). Problems Analysis of Ayandeh Bank branches across the country using data mining method. *Information Technology Management*, 6 (2), 350-333. [In Persian]
- Motameni, Alireza; Moradi, Hadi and Hemmati, Amin. (1391). Identify and rank the factors affecting the loyalty of retailers to distribution companies. *Supply Chain Management Quarterly*, 14 (35), 75-62. . [In Persian]

استناد به این مقاله: رحیمی، فاطمه، سبط، محمد وحید، غنبر تهرانی، نسیم. (۱۴۰۰). تحلیل الگوی رفتاری مشتریان شعب به روش خوشه‌بندی و دسته‌بندی با استفاده از روش RFM، مطالعات مدیریت کسب و کار هوشمند، ۹(۳۶)، ۲۰۹-۲۴۲. DOI: 10.22054/IMS.2021.50853.1697



Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License..

