Sciences and Techniques of
Information Management
Volume 11, Issue 4, 2025
https://stim.qom.ac.ir

# Proposing a New Framework Based on the RFM Model and Multivariate Time Series for Customer Segmentation and Behavior Analysis: A Case Study of a Food Industry Company

**Reza Molaei[1], Hossein Abbasimehr[2] and Kheirollah Rahsepar Fard[3]✉**

1. Department of Computer and Information Technology, Faculty of Engineering, University of Qom, Qom, Iran. Email: r.molaei@stu.qom.ac.ir
2. Faculty of Information Technology and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran. Email: abbasimehr@azaruniv.ac.ir
3. Department of Computer and Information Technology, Faculty of Engineering, University of Qom, Qom, Iran. Email: rahsepar@qom.ac.ir

**Article Info**

**Article type:**
Research Article

**Article history:**
Received: 2025/08/23
Received in revised form: 2025/09/28
Accepted: 2025/11/09
Published online: 2026/01/28

**Keywords:**
Customer segmentation, RFM model, Multivariate time series, Ensemble clustering, Shapley additive explanations, TOPSIS.

**ABSTRACT**

**Purpose:** The objective of this study is to analyze customer behavior using multivariate time-series data. After conducting a comprehensive analysis of customer behavior, customers are prioritized using the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) multi-criteria decision-making method. The results of this study can assist organizations in developing more effective and targeted marketing strategies.

**Method:** In this research, customers' dynamic behavior is analyzed using the RFM (recency, frequency, monetary value) model represented as multivariate time-series data. This approach is considered one of the most recent and practical methods for analyzing customer behavior over time. First, an ensemble time-series clustering method is applied to identify customer clusters and analyze their behavioral patterns from different perspectives. Subsequently, key features are extracted from each time series and used as inputs to a classification model. Finally, the classifier model is interpreted using Shapley Additive Explanations (SHAP), which enables the calculation of the importance of each key feature. These key features and their corresponding weights are then used as inputs for the TOPSIS multi-criteria decision-making method to prioritize customers.

**Findings:** The findings indicate that representing customer purchase data as multivariate time-series data based on RFM variables enables the identification of groups of customers with similar behavioral patterns over time. The proposed approach simultaneously considers RFM variables across time and provides a dynamic analysis of customer behavior. In addition, the application of the SHAP method allows for the interpretation of the classification model and the determination of the importance of critical behavioral features. Based on these weighted features, the TOPSIS method is applied to rank customers according to their importance and priority within marketing strategies. These results can support organizations in designing targeted and effective marketing initiatives.

**Conclusion:** The results demonstrate that the proposed approach enables the identification of customer behavioral patterns and facilitates more accurate customer prioritization. These analyses help organizations better understand customer dynamics and focus their marketing strategies on high-value customers. Overall, the findings support organizations in improving marketing effectiveness and efficiency through a deeper understanding and analysis of customer behavior patterns.

**Cite this article:** Ghaeimi, F., & et al. (2025)., Transforming Education with Artificial Intelligence: A Systematic Review of Applications, Potentials, and Outcomes. *Science and Technology of Information Management*, 11 (4). 56-76. **DOI**: https://doi.org/10.22091/STIM.2025.11525.2174

© The Author(s)
DOI: 10.22091/STIM.2025.11525.2174

**Publisher**: University of Qom

علوم و فنون مدیریت اطلاعات

دوره ۱۱، شماره ۴، ۱۴۰۴

https://stim.qom.ac.ir

# ارائهٔ چارچوبی جدید مبتنی‌بر مدل RFM و سری‌های زمانی چندمتغیره برای بخش‌بندی و تحلیل رفتار مشتریان: مطالعهٔ موردی یک شرکت صنایع غذایی

رضا مولایی[۱]، حسین عباسی مهر[۲] و خیراله رهسپار فرد[۳]

۱. دکتری، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه قم، قم، ایران. رایانامه: r.molaei@stu.qom.ac.ir

۲. دانشیار، دانشکده فناوری اطلاعات و مهندسی کامپیوتر، دانشگاه شهید مدنی آذربایجان، تبریز، ایران. رایانامه: abbasimehr@azaruniv.ac.ir

۳. استادیار، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه قم، قم، ایران(نویسندهٔ مسئول). رایانامه: rahsepar@qom.ac.ir

**اطلاعات مقاله**

**نوع مقاله:** مقاله پژوهشی

**تاریخ دریافت:** ۱۴۰۴/۰۶/۰۱

**تاریخ بازنگری:** ۱۴۰۴/۰۷/۰۶

**تاریخ پذیرش:** ۱۴۰۴/۰۸/۱۸

**تاریخ انتشار:** ۱۴۰۴/۱۱/۰۸

**کلیدواژه‌ها:**

بخش‌بندی مشتریان، مدل RFM، سری زمانی چند متغیره، خوشه‌بندی ترکیبی، روش توضیح افزودنی شاپلی، تاپسیس.

**چکیده**

**هدف:** هدف از این پژوهش تحلیل رفتار مشتریان با استفاده از سری زمانی چند متغیره است. پس از تحلیل و شناخت دقیق مشتریان، اولویت‌بندی آن‌ها به کمک روش تصمیم‌گیری چند معیاره تاپسیس انجام شده است. نتایج این پژوهش می‌تواند یاری‌رسان سازمان‌ها در تدوین استراتژی بازاریابی باشد.

**روش:** در این پژوهش از رویکرد تحلیل رفتار پویای مشتریان با استفاده از متغیرهای تازگی تعداد و میزان خرید در قالب سری زمانی‌های چندمتغیره که از جدیدترین و کاربردی‌ترین روش‌های تحلیل رفتار مشتریان در طول زمان به شمار می‌رود، استفاده شده است. سپس با استفاده از یک روش خوشه‌بندی ترکیبی سری زمانی خوشه‌های مشتریان شناسایی شده، الگوهای رفتاری آنان در ابعاد مختلف شد. پس از این مرحله، ویژگی‌های کلیدی از هر سری زمانی استخراج شده و به‌عنوان ورودی یک مدل طبقه بندی در نظر گرفته شدند. در نهایت با استفاده از روش توضیح افزودنی شاپلی مدل طبقه‌بندی‌کننده بررسی شده و اهمیت هر کدام از ویژگی‌های کلیدی محاسبه می شود. ویژگی‌های کلیدی به‌همراه وزن‌های آن‌ها در روش تصمیم‌گیری چند معیاره تاپسیس برای اولویت‌بندی مشتریان استفاده می‌شود.

**یافته‌ها:** یافته‌های این پژوهش نشان می‌دهد که با بازنمایی داده‌های خرید مشتریان در قالب سری زمانی چندمتغیره متشکل از متغیرهای آر.اف.ام امکان شناسایی گروه‌های مشابه مشتریان با الگوهای رفتاری شبیه به هم در طول زمان به وجود می‌آید. رویکرد ارائه‌شده به‌طور هم‌زمان متغیرهای آر.اف.ام را در طول زمان در نظر گرفته و رفتار پویای مشتریان را تحلیل می‌کند. همچنین کاربرد روش توضیح افزودنی شاپلی در محاسبه اهمیت ویژگی‌های کلیدی رفتار مشتری در این پژوهش نشان داده شد. سپس با استفاده از روش تصمیم‌گیری چندمعیارهٔ تاپسیس، مشتریان براساس اهمیت و اولویت در استراتژی بازاریابی رتبه‌بندی شدند. این نتایج می‌تواند به سازمان در تدوین استراتژی‌های بازاریابی هدفمند و مؤثر کمک کند.

**نتیجه‌گیری:** نتایج نشان می‌دهند که رویکرد پیشنهادی پژوهش امکان شناسایی الگوهای رفتاری مشتریان را فراهم می‌کند. این تحلیل‌ها به سازمان کمک می‌کنند تا الگوهای رفتاری مشتریان را به‌طور مؤثرتر شناسایی کرده و با توجه به اولویت‌بندی انجام‌شده، مشتریان ارزشمند را در استراتژی بازاریابی هدف قرار دهد. به‌طورکلی، نتایج این پژوهش به سازمان‌ها کمک می‌کند تا با شناخت و تحلیل الگوی رفتاری مشتریان، استراتژی بازاریابی مؤثری را تدوین کرده و بهره‌وری بازاریابی خود را افزایش دهند.

**استناد:** مولایی، رضا و دیگران. (۱۴۰۴). «ارائهٔ چارچوبی جدید مبتنی‌بر مدل RFM و سری‌های زمانی چندمتغیره برای بخش‌بندی و تحلیل رفتار مشتریان: مطالعهٔ موردی یک شرکت صنایع غذایی». *علوم و فنون مدیریت اطلاعات*. دورهٔ ۱۱. شمارهٔ ۴. صص: ۷۶-۵۶.
https://doi.org/10.22091/STIM.2025.11525.2174

ناشر: دانشگاه قم

© نویسندگان.

CC BY NC

## ۱. مقدّمه

برای رسیدن به بازاریابی مؤثر در عصر رقابتی حاضر، سازمان‌ها نیاز دارند تا رفتار مشتریان خود را تحلیل و ارتباط مؤثری با آن‌ها را برقرار کنند (ولایتی محمد و دیگران، ۱۳۹۷). در سال‌های اخیر، پژوهش‌ها در حوزهٔ مدیریت ارتباط با مشتری در صنایع با بازار رقابتی بالا توجه زیادی را به خود جلب کرده است (عباسی مهر و شبانی[۱]، ۲۰۲۱ الف؛ گورولا-ناوارو[۲] و دیگران، ۲۰۲۰). با افزایش حجم اطلاعات مشتریان و پیشرفت‌های انجام‌شده در حوزهٔ علم اطلاعات، تقسیم‌بندی مشتریان به یکی از موضوعات داغ پژوهشی در مدیریت ارتباط با مشتری تبدیل شده است که از تکنیک‌های مختلفی مانند تحلیلی آماری داده‌ها، یادگیری عمیق، و الگوریتم‌های یادگیری ماشین به‌شیوهٔ گسترده استفاده می‌شود (باروق[۳] و دیگران ، ۲۰۲۳؛ پاکزاد[۴] و دیگران ، ۲۰۲۳؛ رمضانی[۵] و دیگران ، ۲۰۲۳).

مدل آر.اف.ام[۶] یک تکنیک ساده و مؤثر است که با استفاده از تحلیل کمّی داده‌های مربوط به تراکنش‌ها، برای تقسیم‌بندی مشتریان و تحلیل رفتار آن‌ها استفاده می‌شود (شریفی و دیگران، ۱۴۰۲؛ یوسفی زاد و ثریایی، ۱۳۹۸). در این مدل پارامترهای تازگی (R)، تکرار (F)، و ارزش مالی (M) ابعاد رفتار مشتری به شمار می‌روند. R به معنای مدت زمانی است که از آخرین خرید گذشته است، F تعداد خریدها در یک بازهٔ زمانی را نشان می‌دهد و M نشان‌دهندهٔ مجموع ارزش تمام تراکنش‌ها در یک دورهٔ زمانی است (هاگز[۷]، ۲۰۰۵). مشتریانی که به‌طور هم‌زمان در هر سه بُعد R، F، و M امتیاز بالایی داشته باشند، مشتریان باارزش به شمار می‌روند (هو[۸] و دیگران ،۲۰۱۳ ). در بسیاری از مطالعات پیشین از مدل RFM برای تقسیم‌بندی مشتریان استفاده شده است (آخوندزاده‌نوقابی و البدوی[۹]، ۲۰۱۵) و این مدل یک تکنیک پُرکاربرد برای تحلیل رفتار ارزش مشتری محسوب می‌شود (هلدت[۱۰] و دیگران ، ۲۰۲۱). هدف نهایی شناسایی و ارزش‌گذاری مشتریان هدف برای داشتن استراتژی بازاریابی مؤثر است و بر این اصل که ۸۰ درصد از بهره‌وری کسب‌وکار شما از ۲۰ درصد از مشتریان حاصل می‌شود، استوار است (السیات[۱۱]، ۲۰۲۳).

چالش اصلی در استفاده از مدل RFM سنتی این است که اغلب رویکرد ایستا برای تقسیم‌بندی مشتریان استفاده شده است (عباسی‌مهر و شیخ‌باقری، ۲۰۲۲؛ وانگ[۱۲] و دیگران ، ۲۰۲۴؛ وی[۱۳] و دیگران ، ۲۰۱۹). در مدل‌های تقسیم‌بندی ایستا، مشتری و عملکرد آن ثابت در نظر گرفته می‌شود، درحالی‌که در واقعیت رفتار مشتری در طول زمان تغییر می‌کند و تکامل می‌یابد (عباسی مهر و بحرینی[۱۴]، ۲۰۲۲؛ عباسی‌مهر و شبانی، ۲۰۲۱ الف؛ آخوندزاده‌نوقابی و البدوی، ۲۰۱۵). در سال‌های اخیر تمرکز اصلی پژوهشگران بر این موضوع بوده است که تحلیل رفتار مشتریان در یک زمان خاص و به‌صورت ایستا کفایت نمی‌کند و باید روندهای رفتاری مشتریان در طول زمان بررسی شود. بر این اساس، تقسیم‌بندی پویای مشتریان از اهمیت بالایی برخوردار است. اخیراً، ترکیب مدل RFM و خوشه‌بندی سری‌های زمانی به‌عنوان روشی برای تقسیم‌بندی پویای مشتریان و تحلیل رفتار مطرح شده است (عباسی‌مهر و دیگران ، ۲۰۲۲؛ عبادی جلال و المغربی[۱۵]، ۲۰۲۴؛ وانگ و دیگران ، ۲۰۲۴).

خوشه‌بندی سری‌های زمانی به سه دستهٔ اصلی تقسیم می‌شود: روش‌های مبتنی‌بر فاصله، مبتنی‌بر ویژگی، و مبتنی مدل. در روش‌های مبتنی‌بر فاصله، فاصلهٔ جفتی میان سری‌های زمانی با استفاده از یک معیار فاصله محاسبه می شود. انتخاب یک معیار فاصلهٔ مناسب چالش‌برانگیز است و باید عوامل مختلفی ازجمله نویز، طول‌های متفاوت سری‌ها، و دیگر مقیاس‌های مختلف لحاظ

---

1. Abbasimehr & Shabani
2. Guerola-Navarro
3. Barough
4. Pakzad
5. Ramezani
6. Recency, Frequency, and Monetary (RFM)
7. Hughes
8. Hu
9. Akhondzadeh-Noughabi & Albadvi
10. Heldt
11. Alsayat
12. Wang
13. Wei
14. Abbasimehr, & Bahrini
15. Ebadi Jalal & Elmaghraby

شـود (آقا بزرگی[1] و دیگران ، ۲۰۱۵). ترکیب مدل آر.اف.ام و تحلیل سـری‌های زمانی اغلب می‌تواند تحلیل دقیقی از ارزش مشـتری ارائه دهد و تحلیل رفتار مشتریان در طول زمان را میسر سازد (عباسی‌مهر و شبانی، ۲۰۲۱ الف). بااین‌حال، این رویکرد با چالش‌هایی همراه اسـت. یکی از این چالش‌ها این اسـت که در بیشـتر مطالعات، تمرکز بر سـری‌های زمانی تک متغیره بوده و تنها بر بُعد میزان خرید تمرکز شده است و دیگر فاکتورهای مهم در رفتار مشتریان در نظر گرفته نشده است. علاوه‌بر این چالش‌ها، در مطالعات قبلی از روش‌های منفرد خوشه‌بندی سری زمانی استفاده شده است، در حالی‌که استفاده از یک رویکرد خوشه‌بندی ترکیبی[2] امکان استفاده از مزایای روش‌های متنوع خوشه‌بندی سری زمانی را ممکن می‌سازد.

همچنین، در اکثر پژوهش‌ها، خروجی و ترکیب این مدل‌ها به‌طور عملی در اختیار مدیران سازمانی قرار نمی‌گیرد. این مو ضوع باعث می‌شود که مدیران نتوانند از این تحلیل‌ها برای شناسایی مشتریان باارزش استفاده کنند و درنتیجه، استراتژی‌های بازاریابی خود را بهینه سازی کنند. درواقع، نبودِ د سترسی به رتبه‌بندی مشتریان در تصمیم‌گیری‌های بازاریابی می‌تواند به عدم شنا سایی مشتریان کلیدی منجر شود که نقش مهمی در رشد و توسعهٔ سازمان ایفا می‌کنند.

با توجه به چالش‌های مرتبط با ترکیب مدل آر.اف.ام و تحلیل سـری‌های زمانی، در مطالعهٔ حاضـر یک رویکرد پویا برای تقسـیم‌بندی مشـتریان ارائه شـده اسـت. رویکرد ارائه‌شـده بر ترکیب مدل آر.اف.ام و تحلیل سـری‌های زمانی چندمتغیره تمرکز دارد. ویژگی‌های تکرار خرید و مجموع ارزش تراکنش‌ها نیز متغیرهای اصـلی این تحلیل به شـمار می‌روند. روش پیشـنهادی با اسـتفاده از داده‌های مشتریان شرکتی[3] در یک شرکت صنایع غذایی واقعی ارزیابی شده است. این روش که به نام تقسیم‌بندی پویا و رتبه‌بندی مشتریان شناخته می‌شود، به شرکت‌های صنایع غذایی کمک می‌کند تا مشتریان ارزشمند خود را شناسایی و مدیریت کنند. در ابتدا، لیسـت مشـتریان فعال با اسـتفاده از مراحل پیش‌پردازش داده‌ها، به دسـت می‌آید. سـپس، ویژگی‌های تکرار خرید و مجموع ارزش تراکنش‌ها از داده‌های مشتریان استخراج می‌شود. برای خوشه‌بندی مشتریان، از یک رویکرد خوشه‌بندی ترکیبی استفاده می‌شود. این روش با ترکیب چندین الگوریتم خوشـه‌بندی به بهبود کیفیت خوشـه‌بندی کمک می‌کند و نسـبت به روش‌های سـنتی، دقت و پایداری نتایج بهتری دارد (بونگون و یم-آن[4]، ۲۰۱۸). به‌عبارتی دیگر، رویکرد خوشـه‌بندی ترکیبی به ما این امکان را می‌دهد که از نقاط قوت هر الگوریتم بهره‌مند شـویم و نتایج دقیق‌تری به دست آوریم. پس از انجام خوشـه‌بندی، مجموعه‌ای از ویژگی‌های کلیدی هر سری زمانی که در مطالعهٔ هیندمن[5] و دیگران(۲۰۱۵) پیشنهادشده است، استخراج می‌شود. برای نشان دادن شاخص‌های مؤثر در فرایند خوشه‌بندی، از روش شاپلی[6] (لوندبرگ[7]، ۲۰۱۷) استفاده شده است. پس از شناسایی و تعیین این شاخص‌ها، با بهره‌گیری از الگوریتم تصمیم‌گیری چندمعیاره[8] تاپسیس[9]، یک رتبه‌بندی مناسب از مشتریان براساس استراتژی‌های بازاریابی مؤثر ارائه خواهد شد. این رویکرد می‌تواند به مدیران سـازمانی در شـناسـایی و هدف‌گذاری مشـتریان کلیدی در محیط‌های واقعی کمک کند و درنهایت به بهینه‌سازی استراتژی‌های بازاریابی منجر شود.

به‌طور خلاصه نوآوری‌های اصلی این پژوهش به شرح زیر است:

- این مطالعه یک رویکرد پویا برای تقسـیم‌بندی مشـتریان با تأکید بر ترکیب مدل آر.اف.ام و تحلیل سـری‌های زمانی چندمتغیره براسـاس ویژگی‌های تکرار خرید (F) و مجموع ارزش تراکنش‌ها (M) ارائه می‌دهد که به شـرکت‌های صنایع غذایی در شناسایی و مدیریت مشتریان عمده کمک می‌کند.
- برای خوشه‌بندی مشتریان، از رویکرد خوشه‌بندی ترکیبی استفاده می‌شود که با ترکیب چندین الگوریتم، دقت و پایداری نتایج را افزایش می‌دهد و به درک بهتر رفتار مشتریان کمک می‌کند.

1. Aghabozorgi
2. Ensemble clustering
3. Business customer
4. Boongoen & Iam-On
5. Hyndman
6. SHapley Additive exPlanations (SHAP)
7. Lundberg
8. Multiple Criteria Decision Makin (MCSM)
9. Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

- ویژگی‌های کلیدی رفتار مشتری از سری‌های زمانی استخراج شده و با استفاده از روش توضیح افزودنی مبتنی‌بر شاپ، اهمیت ویژگی‌های کلیدی محاسبه می‌شود.
- با بهره‌گیری از الگوریتم تصمیم‌گیری چندمعیارهٔ تاپسیس، یک رتبه‌بندی مناسب از مشتریان بر اساس استراتژی‌های بازاریابی مؤثر ارائه می‌شود که مدیران سازمانی را در شناسایی و هدف‌گذاری مشتریان کلیدی در محیط‌های واقعی یاری می‌کند.

ادامه مقاله به شرح زیر است: در بخش دوم، مفاهیم و کارهای پیشین بررسی می‌شوند. در بخش سوم، مدل پژوهش به تفصیل معرفی می‌شود. بخش چهارم به نتایج و بحث اختصاص دارد و درنهایت، در بخش پایانی به جمع‌بندی و پیشنهادهایی برای پژوهش‌های آینده پرداخته می‌شود.

## ۲. مبانی نظری و کارهای پیشین

در دنیای رقابتی امروز، بخش‌بندی پویای مشتریان و تصمیم‌گیری مبتنی‌بر داده‌ها دو مؤلفه کلیدی در موفقیت کسب‌وکارها هستند. با توجه به تغییرات مداوم در رفتار مشتریان و پیچیدگی‌های بازار، رویکردهای سنتی بخش‌بندی مشتریان دیگر قادر به برآورده کردن نیازهای سازمان‌ها نیستند. به همین دلیل، پژوهش در حوزهٔ بخش‌بندی پویای مشتریان و استفاده از آن برای تصمیم‌گیری مؤثر به‌طور فزاینده‌ای در حال افزایش است. این بخش از مقاله به بررسی مبانی نظری و مرور کارهای پیشین می‌پردازد.

### ۲-۱. بخش‌بندی پویای مشتریان

بخش‌بندی مشتریان یکی از ابزارهای اساسی در مدیریت ارتباط با مشتری است که به شرکت‌ها این امکان را می‌دهد تا با شناسایی و طبقه‌بندی مشتریان براساس ویژگی‌های منحصربه‌فرد، استراتژی‌های بازاریابی و خدمات خود را بهینه‌سازی کنند. این فرایند معمولاً با تحلیل رفتار مشتری انجام می‌شود و به دودستهٔ تقسیم‌بندی ایستا و پویا تقسیم می‌شود (آخوندزاده‌نوقابی و البدوی، ۲۰۱۵). نقطه ضعف اصلی رویکردهای ایستا این است که تنها در یک زمان مشخص معتبر هستند و تغییرات رفتاری مشتریان را در طول زمان نادیده می‌گیرند (آخوندزاده‌نوقابی و البدوی، ۲۰۱۵؛ مصدق[1] و دیگران، ۲۰۲۱). با ظهور فناوری‌های جدید و دسترسی به داده‌های بزرگ، روش‌های مبتنی‌بر تحلیل داده و یادگیری ماشین برای بهبود دقت و کارایی تقسیم‌بندی مشتریان به کار گرفته شده‌اند و بهره‌گیری از تحلیل‌های چندمعیاره و هوش تجاری به بهینه‌سازی فرایند تقسیم‌بندی و شناسایی الگوهای رفتاری کمک کرده است (کاتراگادا[2]، ۲۰۲۲؛ مصدق و دیگران، ۲۰۲۱). درنتیجه، رویکردهای پویا که تغییرات رفتاری مشتریان را در طول زمان لحاظ می‌کنند، به کسب‌وکارها در درک بهتر نیازهای مشتری و برقراری ارتباطات مؤثرتر کمک می‌کنند.

موضوع تقسیم‌بندی پویا در حوزه‌های مختلفی بررسی شده است. برای نمونه، در پژوهشی که ونگ[3] و دیگران (۲۰۲۴) انجام داده‌اند، نیاز نمایندگان قطعات خودرو برای تحلیل و مدیریت مؤثر مشتریان در عصر اینترنت بررسی و تقسیم‌بندی پویای مشتریان به‌مثابه ابزار تحلیلی مؤثر معرفی شده است. در این مقاله رویکردی جدید با ترکیب مدل طولعمر مشتری- تازگی- تعداد- میزان و میزان رضایت[4] و خوشه‌بندی سری زمانی ارائه شده است که رفتار مشتریان را به‌صورت دنباله‌ای از متغیرهای کلیدی نمایش می‌دهد. در مقالهٔ عباسی مهر و شیخ‌باقری (۲۰۲۲)، یک روش جدید استخراج ویژگی برای خوشه‌بندی سری زمانی مشتریان معرفی شده است که با استفاده از روش لاپلاسین[5] رتبه‌بندی ویژگی‌ها را انجام می‌دهد و در مقایسه با بقیهٔ روش‌ها عملکرد بهتری دارد. همچنین، یک مدل پیش‌بینی رفتار مشتری به‌نام مدل ماشین بردار پشتیبان بهبود یافته با الگوریتم بهینه سازی ملخ ارائه شده است که در اکثر موارد ازنظر دقت پیش‌بینی از سایر مدل‌ها بهتر عمل می‌کند. در پژوهش دیگری که اخیراً انجام شده است (عبادی جلال و المغربی، ۲۰۲۴)، محدودیت‌های پژوهش‌های پیشین در بخش‌بندی پویای مشتریان بررسی شده و یک چارچوب کارآمد برای پیش‌بینی رفتار مشتریان و تحلیل انتقال آن‌ها بین گروه‌ها ارائه شده است. در مقالهٔ عبادی جلال و المغربی از الگوریتم K-means برای خوشه‌بندی استفاده شده است، که چالش‌هایی مطرح است. زیرا این الگوریتم

1. Mosaddegh
2. Katragadda
3. Wang
4. Length, Recency, Frequency, and Monetary, Satisfaction (LRFMS)
5. Laplacian

خوشه‌بندی ممکن است به دلیل بی‌توجهی به توزیع و ساختار زمانی داده‌ها، دقت کمتری داشته باشد. همچنین، لحاظ کردن معیارهای فاصله‌ای خاص مانند فاصلهٔ اقلیدسی و منهتن ممکن است به ندا شتنِ دقت در تحلیل رفتار مشتریان منجر شود. در ادامه، یک سری از مطالعات پژوهشی پیشین در زمینهٔ تقسیم‌بندی مشتریان در جدول ۱ خلاصه شده است.

**جدول ۱. خلاصه‌ای از کارهای پیشین**

| مقاله | حوزهٔ پژوهش | الگوریتم/ تکنیک استفاده‌شده | خروجی پژوهش |
|---|---|---|---|
| (خواجه‌وند و تارخ[1]، ۲۰۱۱) | بانکداری | K-means, two-step, x-means, autoregressive integrated moving average (ARIMA), AHP | مشتریان ابتدا براساس متغیرهای آر.اف.ام گروه‌بندی و سپس با استفاده از آر.اف.ام وزن‌دار و روش ای.اچ.پی رتبه‌بندی شدند. برای پیش‌بینی ارزش طول عمر مشتری از الگوریتم ای.آر.آی.آم.ای استفاده شده است. |
| (پروانه[2] و همکاران، ۲۰۱۲) | شرکت فعال در صنعت بهداشتی | Weighted LRFM (Length RFM), K-means, AHP | این مطالعه با ترکیب ای.اچ.پی و ال.آر.اف.ام، خرده‌فروشان صنعت بهداشتی را به شش گروه تقسیم کرده و به بهبود استراتژی‌های بازاریابی و روابط تولیدکننده و خرده‌فروشان کمک می‌کند. |
| (کوسمنت[3] و همکاران، ۲۰۱۴) | بازاریابی | RFM, CHAID decision tree and logistic regression | این مقاله نشان می‌دهد که در حالت کلی برای بخش‌بندی مشتریان برای بازاریابی مستقیم، درخت تصمیم از تحلیل آر.اف.ام و رگرسیون لجستیک برتر است و به‌عنوان گزینه‌ای مناسب در شرایط نادرستی داده‌ها توصیه می‌شود. |
| (سانگ[4] و همکاران، ۲۰۱۶) | خدمات تلفن همراه | RFM, Time Series Analysis, MCA (Multiple Correspondence Analysis) | این مقاله مدل آر.اف.ام را با یک رویکرد آماری جدید بهبود می‌بخشد و با تحلیل داده‌های بزرگ در خدمات تلفن همراه، به چالش‌های سنتی سی.آر.ام پاسخ می‌دهد و روش‌های بهینه‌سازی برای بخش‌بندی کاربران را معرفی می‌کند. |
| (دوغان[5] و همکاران، ۲۰۱۸) | خرده‌فروشی ورزشی | Two-step cluster analysis, K-means | استفاده از تحلیل آر.اف.ام و تکنیک‌های خوشه‌بندی برای انجام بخش‌بندی مشتریان فروشگاه خرده‌فروشی ورزشی و درنتیجه طراحی استراتژی‌های بازاریابی مناسب. |
| ((عباسی مهر و شبانی، ۲۰۱۹) | تراکنش مشتریان بانک | RFM, time series | در این مقاله روشی برای پیش‌بینی رفتار آینده و بخش‌بندی مشتریان براساس داده‌های آن‌ها ارائه شده است. با استفاده از تکنیک‌های خوشه‌بندی و سری‌های زمانی، روشی با دقت بالاتر برای تحلیل رفتار مشتریان در بازاریابی سازمان‌ها ارائه شده است. |
| (عباسی مهر و شبانی، ۲۰۲۱ ب) | تراکنش مشتریان بانک | RFM, time series | ایجاد یک چارچوب برای نمایش رفتار مشتریان با استفاده از سری‌های زمانی و سپس به‌کارگیری روش خوشه‌بندی سری‌های زمانی مبتنی‌بر فاصله برای گروه‌بندی مشتریان |
| (پریامباداریا[6] و همکاران، ۲۰۲۱) | فرآیندهای آموزشی | Cluster Evolution Analysis, Profile-based Cluster Evolution Analysis, K-means | روش جدید برای تحلیل رفتار یادگیری دانش‌آموزان با استفاده از تحلیل تکامل خوشه‌ای ارائه می‌دهد و به ذی‌نفعان آموزشی در بهبود برنامه‌ریزی درسی و پیشگیری از ترک تحصیل کمک می‌کند. |

1. Khajvand & Tarokh
2. Parvaneh
3. Coussement
4. Song
5. Doğan
6. Priyambada

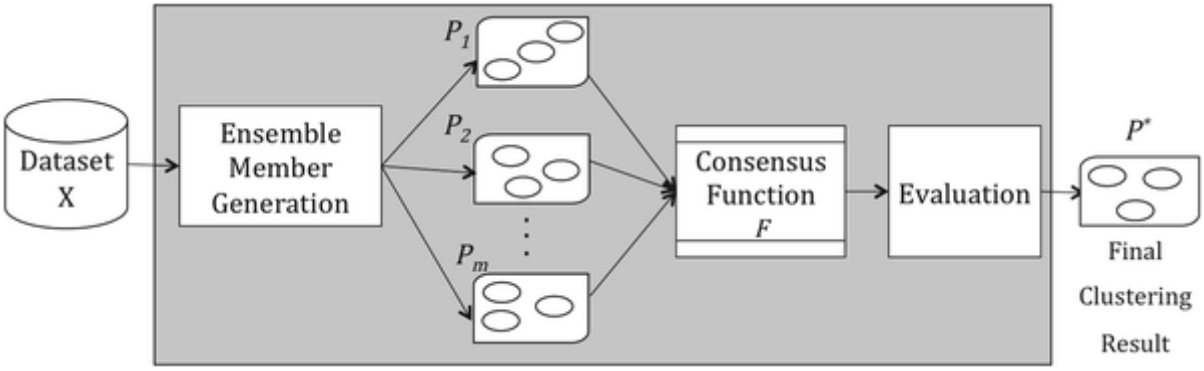
| مقاله | حوزهٔ پژوهش | الگوریتم/ تکنیک استفاده‌شده | خروجی پژوهش |
|---|---|---|---|
| (مصدق و همکاران، ۲۰۲۱) | تجزیه‌وتحلیل داده‌های بزرگ | CLV (Customer Lifetime Value), Pattern Mining, | ارائهٔ یک روش برای دسته‌بندی مشتریان براساس ارزش عمر مشتری و ردیابی انتقال مشتریان در بین دسته‌ها با بهره‌گیری از استخراج قوانین ارتباطی. |
| (عباسی مهر و شیخ‌باقری، ۲۰۲۲) | تراکنش‌های نقطه فروش POS | Feature-based Time Series Clustering, Support Vector Regression | ارائهٔ یک تکنیک نوین برای خوشه‌بندی مشتریان براساس سری‌های زمانی و ویژگی‌ها ارائه می‌دهد که به بهبود تقسیم‌بندی و پیش‌بینی رفتار مشتریان کمک می‌کند. همچنین، رگرسیون بردار پشتیبان به‌مثابه ابزاری مؤثر برای پیش‌بینی معرفی می‌شود. |
| (لیو و چن[1]، ۲۰۲۲) | مدیریت داده‌های بهداشتی در بستر IOT | RFM, Density-based Spatial Clustering, Grouping Genetic Algorithm | این مقاله مدل آر.اف.ام را با الگوریتم‌های ژنتیک گروهی و دی.بی.اسکن بهبود می‌بخشد و دقت در انتخاب ویژگی و خوشه‌بندی را بهبود می‌دهد. |
| (اسمعیلی و حکیمی[2]، ۲۰۲۳) | خرده‌فروشی | K-Means, RFM-D (Recency, Frequency, Monetary, and Diversity) | این مقاله مدل آر.اف.ام-دی را معرفی می‌کند که با افزودن تنوع به تقسیم‌بندی مشتریان، دقت پیش‌بینی رفتار آن‌ها را افزایش می‌دهد و به شرکت‌ها در هدف‌گذاری مؤثرتر کمک می‌کند. |
| (سان[3] و همکاران، ۲۰۲۳) | خرده‌فروشی آنلاین | BG/NBD (Beta-Geometric/Negative Binomial Distribution model) | این مقاله روشی جدید برای اندازه‌گیری ارزش مشتری در روابط غیر قراردادی ارائه می‌دهد که با داده‌های واقعی تأییدشده و دقت تحلیل و استراتژی‌های بازاریابی را بهبود می‌بخشد. |
| (عبادی جلال و المغربی، ۲۰۲۴) | مشتریان یک شرکت فناوری اطلاعات | K-means, time series, Counterfactual Analysis | این مقاله چارچوبی برای پیش‌بینی رفتار مشتریان و تحلیل انتقالات آن‌ها ارائه می‌دهد و به تصمیم‌گیرندگان در توسعهٔ استراتژی‌های بازاریابی شخصی‌سازی‌شده کمک می‌کند. |
| (ونگ و همکاران، ۲۰۲۴) | مشتریان قطعات خودرو | LRFMS (Length, Recency, Frequency, Monetary, and Satisfaction), time series | این مقاله روشی نوین برای بخش‌بندی مشتریان در صنعت قطعات خودرو ارائه می‌دهد که با ترکیب مدل ال.آر.اف.ام.اس و خوشه‌بندی سری زمانی، به شناسایی رفتار مشتریان و پیشنهاد استراتژی‌های بازاریابی مؤثر می‌پردازد. |
| (دانوشکدی[4] و همکاران، ۲۰۲۴) | داده‌های خرید مشتریان | RFM, Neural network classifiers, SVM | این مقاله روشی برای بخش‌بندی، تحلیل و پیش‌بینی رفتار مشتریان در صنعت خرده‌فروشی سوپرمارکت ارائه می‌دهد. این روش شامل مراحلی ازجمله پیش‌پردازش داده‌ها، تحلیل اکتشافی، مهندسی ویژگی‌ها، انتخاب مدل، و ارزیابی است. |
| (پیتر[5] وهمکاران، ۲۰۲۵) | داده‌های کمپین‌های بازاریابی تلفنی بانک‌ها | Ensemble learning models (bagging, boosting, and stacking) | در این مطالعه یک مدل یادگیری ترکیبی برای افزایش دقت مدل پیش‌بینی اشتراک‌های مشتریان در کمپین‌های بازاریابی تلفنی بانک‌ها ارائه شده است. در این مدل متغیرهای کلیدی مانند مدت |

1. Liu & Chen
2. Smaili & Hachimi
3. Sun
4. Dhanushkodi
5. Peter

| مقاله | حوزهٔ پژوهش | الگوریتم/ تکنیک استفاده‌شده | خروجی پژوهش |
|---|---|---|---|
| | | | زمان تماس، شاخص‌های اقتصادی، و سن مشتریان در پیش‌بینی اشتراک مشتریان در نظر گرفته شده است. |
| روش پیشنهادی مقاله | مشتریان شرکت صنایع غذایی | RFM, multivariate time series, ensemble clustering | در این مقاله روشی مبتنی‌بر سری‌های زمانی چندمتغیره برای تحلیل رفتار پویای مشتریان و شناسایی الگوهای رفتاری آن‌ها با استفاده از متغیرهای آر.اف.امارائه شده است. همچنین اهمیت ویژگی‌ها و اولویت‌بندی مشتریان با استفاده از ترکیب روش شَپ تاپسیس انجام شده است. |

## ۲-۲. مروری بر روش‌های خوشه‌بندی سری‌های زمانی

به‌طورکلی، دو چالش عمده در الگوریتم‌های خوشه‌بندی وجود دارد (بونگون و یم-آن، ۲۰۱۸؛ دودا[1] و دیگران، ۲۰۰۰): نخست، الگوریتم‌های مختلف روی یک مجموعه داده مشابه، ساختارهای متفاوتی (مانند اندازه و شکل خوشه) را کشف می‌کنند. برای مثال، $k$-means که شناخته‌شده‌ترین تکنیک است، برای خوشه‌های با شکل کروی مناسب است، درحالی‌که خوشه‌بندی سلسله‌مراتبی برای شناسایی الگوهای متصل مؤثر است. این تفاوت به این دلیل است که هر الگوریتم به‌طور ویژه برای بهینه‌سازی یک معیار خاص طراحی شده است. دوم، یک الگوریتم خوشه‌بندی واحد با تنظیمات پارامتری مختلف می‌تواند ساختار خوشه‌ای متنوعی را در همان مجموعه داده نشان دهد. یک تنظیم خاص ممکن است برای برخی مجموعه‌ها مناسب باشد، اما برای همه مناسب نیست. این چالش‌ها انتخاب تکنیک مناسب خوشه‌بندی را بسیار دشوار می‌کند. برای رفع این چالش‌ها، روش‌های خوشه‌بندی ترکیبی ارائه شدند که در آن‌ها ابتدا چندین الگوریتم خوشه‌بندی پایه به مجموعه داده $\boldsymbol{X}$ اعمال می‌شود تا نتایج خوشه‌بندی متنوعی $(\mathbf{p_1}, \mathbf{p_2}, \ldots \mathbf{p_m})$ به دست آید. سپس، این خروجی‌ها برای ایجاد نتیجهٔ نهایی خوشه‌بندی $(\boldsymbol{p}*)$ با استفاده از یک تابع اجماعی ترکیب می‌شوند (بونگون و یم-آن، ۲۰۱۸؛ یم-آن و دیگران، ۲۰۱۰). شکل ۱ فرایند خوشه‌بندی ترکیبی را نشان می‌دهد.




**شکل ۱. فرایند پایه‌ای خوشه‌های ترکیبی (القریشی و وانگ[2]، ۲۰۱۹)**

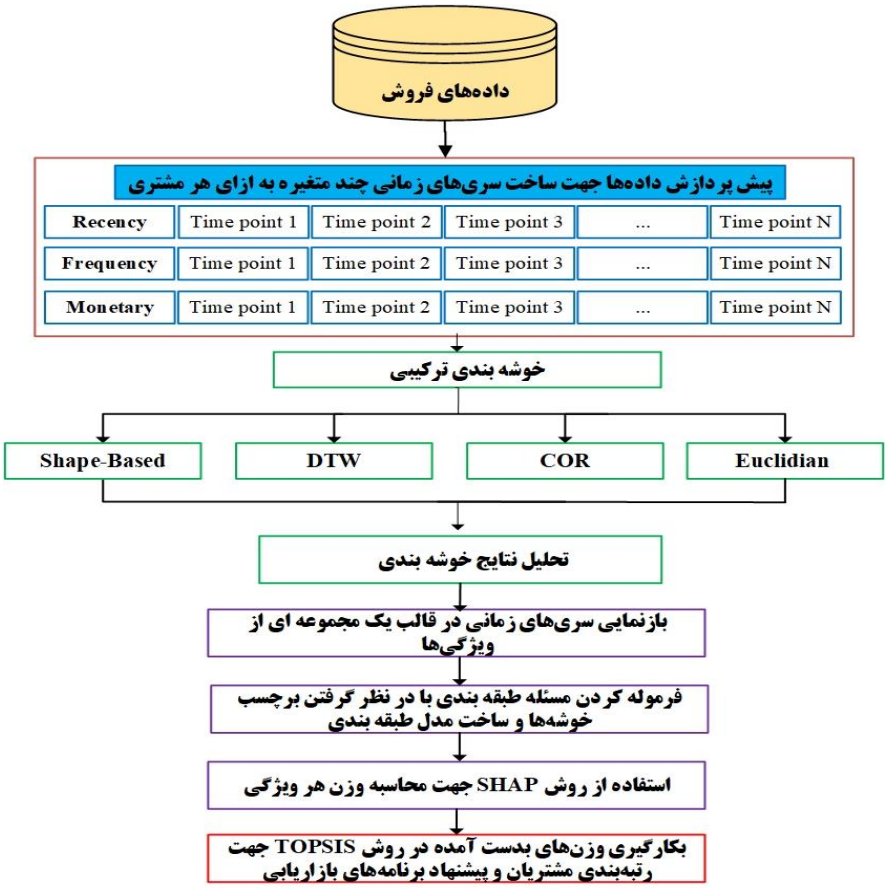
پژوهش‌های زیادی دربارهٔ مفاهیم و یافته‌های جدید در خوشه‌بندی ترکیبی منتشر شده است. برخی از این مطالعات به بهبودهای نظری و گسترش روش‌های قبلی پرداخته و برخی دیگر بر کاربردهای عملی و استفاده از روش‌های موجود در مسائل واقعی تمرکز دارند. برای مثال، خوشه‌بندی ترکیبی در سری‌های زمانی به شناسایی الگوهای مختلف و بهبود دقت تحلیل کمک

1. Duda
2. Alqurashi & Wang

می‌کند که در پژوهش‌هایی که اخیراً انجام شده‌اندبه کاربردهای خوشه‌بندی در تحلیل سری‌های زمانی پرداخته شده است (راماسو[1] و دیگران، ۲۰۱۵؛ یانگ و ژیانگ[2]، ۲۰۱۴). همچنین توجه به شـباهت بین قطعات سـری زمانی و ترکیب آن‌ها به‌منظور بهبود دقت خوشه‌بندی مورد بررسی قرار گرفته است (قربانیان و رضوی، ۱۴۰۳). از دیگر کاربردهای خوشه‌بندی ترکیبی، استفاده از آن در تحلیل شـبکه‌های پیچیده اسـت. در این شـبکه‌ها، خوشـه‌بندی ترکیبی برای ترکیب گروه‌بندی‌های گره‌ها به کار رفته اسـت که شـامل حذف یال‌های ضعیف بین گره‌هایی است که به‌ندرت در یک گروه قرار می‌گیرند (کیم[3] و دیگران، ۲۰۱۲).

## ۳. روش پیشنهادی

روش پیشنهادی برای تحلیل رفتار پویای مشتری در شکل ۲ نشان داده شده است. این چارچوب پیشنهادی به‌طور مؤثری مدل آر.اف.ام را با خوشه‌بندی سری‌های زمانی ترکیب می‌کند. همان‌طورکه در شکل مشخص است این چارچوب شامل ۱) پیش‌پردازش داده‌ها و تشکیل سری‌های زمانی چند متغیره، ۲) خوشه‌بندی سری‌های چند متغیره با استفاده از یک روش خوشه‌بندی ترکیبی، ۳) برچسب‌گذاری و تحلیل رفتاری خوشه‌ها، ۴) استخراج ویژگی‌هایی ازسری‌های زمانی و فرموله کردن مسئلهٔ طبقه‌بندی با در نظرگرفتن برچسب خوشه به‌مثابه متغیر هدف و آموزش یک مدل مبتنی بر روش تقویت گرادیان[4]، ۵) استفاده از روش شَپ برای محاسبهٔ وزن ویژگی‌ها، و ۶) در نظر گرفتن وزن ویژگی‌ها و استفاده از روش تاپسیس برای رتبه‌بندی مشتریان. در ادامه، جزئیات هر مرحله در این چارچوب توصیف شده است.




شکل ۲. چارچوب پیشنهادی

1. Ramasso
2. Yang & Jiang
3. Kim
4. Xgboost

### ۳-۱. پیش‌پردازش داده‌ها و تشکیل سری‌های زمانی چند متغیره

در این مرحله، داده‌های فروش دارای برچسب زمانی به بازه‌های زمانی مشخص (برای مثال، بازه‌های ماهانه) تقسیم می‌شوند. سپس، برای هر مشتری سری زمانی مربوط به هر متغیر مدل آر.اف.ام محاسبه می‌شود. تعریف سری زمانی برای متغیرهای آر.اف.ام به شرح زیر است:

- Recency(t): مدت زمانی که از آخرین خرید مشتری در آن بازه تا پایان بازهٔ انتخاب‌شده سپری شده است.
- Frequency(t): تعداد خریدهایی که در طول بازهٔ زمانی انجام شده است.
- Monetary(t): مقدار تجمیع‌شدهٔ مبلغ خرید که در طول بازهٔ زمانی رخ داده است.

### ۳-۲. خوشه‌بندی مشتریان با استفاده از روش پیشنهادی

یکی از تصمیمات مهم در انجام خوشه‌بندی سری زمانی انتخاب معیار فاصله است. با توجه به اینکه معیارهای فاصلهٔ موجود روی سری‌های تک‌متغیره تعریف شده‌اند در این بخش معیارهای فاصله با توجه به چندمتغیره بودن سری‌های زمانی تعریف شده‌اند:

با فرض داشتن دو سری زمانی چندمتغیره $X$ و $Y$ با ابعاد $n \times d$ که n نشان‌دهنده طول هر سری و d نشان‌دهنده ابعاد هر سری است، در ادامه معیارهای استفاده‌شده برای محاسبهٔ فاصلهٔ دو سری زمانی توصیف می‌شوند.

#### ۳-۲-۱. فاصلهٔ اُقلیدسی[1]

فاصله اقلدیسی بین $X$ و $Y$ با رابطهٔ (۱) تعریف می‌شود که در آن به‌ازای هر کدام از ابعادی سری زمانی، فاصلهٔ دو سری محاسبه می‌شود.

$$d(X,Y) = \sqrt{\sum_{t=1}^{n}\sum_{k=1}^{d}\left(X_t^k - Y_t^k\right)^2} \qquad (1)$$

#### ۳-۲-۲. فاصلهٔ پیچش زمانی پویا[2]

فاصلهٔ پیچش زمانی پویا فاصلهٔ بین دو سری زمانی را با هم‌ترازی آن‌ها ابا استفاده از پیچش محور زمان یک سری به دیگری محاسبه می‌کند (مونترو و ویلار[3]، ۲۰۱۴). روش فاصلهٔ پیچش زمانی پویا برای محاسبهٔ فاصلهٔ بین این دو سری روی هرکدام از ابعاد با استفاده از رابطهٔ (۲) محاسبه می‌شود.

$$DTW_k(X^k, Y^k) = \min_{r\in n}\left(\sum_{t=1}^{n}\left|X^k{}_{it} - Y^k{}_{jt}\right|\right), \qquad (2)$$

در اینجا $DTW_k(X^k, Y^k)$ نشان‌دهندهٔ فاصله فاصلهٔ پیچش زمانی پویا محاسبه شده برای بُعد k –ام سری زمانی است. الگوریتم مسیر پیچش بهینه‌ای را پیدا می‌کند که فاصلهٔ تجمعی بین ابعاد را به حداقل می‌رساند و امکان کشیدگی یا فشردگی محور زمان را فراهم می‌کند در این فرمول، $\boldsymbol{r}$ به‌معنای یافتن مسیر پیچشی بهینه است که مجموع فاصلهٔ تجمعی بین دو سری زمانی $X^k$ و $Y^k$ را به حداقل می‌رساند. هدف یافتن مسیر پیچشی بهینه است که فاصلهٔ کلی بین دو سری را به حداقل می‌رساند (النیمت[4] و دیگران، ۲۰۱۲).

درنهایت فاصله کلی با تجمیع فاصله‌های پیچش زمانی پویای مربوط به هر بُعد با استفاده از رابطهٔ (۲) محاسبه می‌شود.

$$DTW(X,Y) = \sum_{k=1}^{d} DTW_k(X^k, Y^k) \qquad (3)$$

1. Euclidian
2. Dynamic Time Warping (DTW)
3. Montero & Vilar
4. Al-Naymat

### ۳-۲-۳. معیار فاصلهٔ ناوردا پیچیدگی[1]

معیار فاصلهٔ ناوردا پیچیدگی (باتیستا[2] و دیگران، ۲۰۱۴) از اطلاعات مربوط به تفاوت‌های پیچیدگی بین دو سری زمانی به‌عنوان یک عامل تصحیح برای معیارهای فاصلهٔ موجود (مانند فاصله اقلیدسی) استفاده می‌کند. به‌طورکلی این معیار به‌صورت زیر تعریف می‌شود (رابطه (۴)).

$$d_{CID}(X,Y) = CF(X,Y) \cdot d(X,Y) \quad (4)$$

که در آن $d(X,Y)$ به یک معیار فاصله موجود، مانند فاصلهٔ اقلیدسی (رابطه (۱))، اشاره دارد و $CF(X,Y)$ یک عامل تصحیح پیچیدگی است که با رابطهٔ (۵) نشان داده می‌شود:

$$CF(X,Y) = \frac{max(CE(X), CE(Y))}{min(CE(X), CE(Y))} \quad (5)$$

که در آن $CE(X)$ و $CE(Y)$ به‌ترتیب پیچیدگی سری‌های زمانی $X$ و $Y$ را تخمین می‌زنند. برای سری‌های زمانی $X$ و $Y$، $CE(X)$ و $CE(Y)$ به‌ترتیب با روابط (۶) و (۷) محاسبه می شود.

$$CE(X) = \sqrt{\sum_{t=1}^{n}\left(X_t^k - X_{t-1}^k\right)^2} \quad (6)$$

$$CE(Y) = \sqrt{\sum_{t=1}^{n}\left(Y_t^k - Y_{t-1}^k\right)^2} \quad (7)$$

### ۳-۲-۴. فاصلهٔ مبتنی‌بر همبستگی[3]

فاصلهٔ مبتنی‌بر همبستگی (COR) معیاری است که همبستگی بین ابعاد دو سری را برای محاسبهٔ فاصله بین آن‌ها در نظر می‌گیرد. این روش به این نکته توجه دارد که تغییرات یک سری زمانی تا چه حد با تغییرات سری زمانی دیگر مطابقت دارد که آن را به‌ویژه برای سری‌های زمانی که الگوها یا روندهای مشابهی دارند اما ممکن است مقیاس‌ها یا اندازه‌های متفاوتی داشته باشند، مفید می‌سازد (چواکریا و ناگابوشان[4]، ۲۰۰۷؛ مونترو و ویلار، ۲۰۱۵).

معیار COR برای دو سری زمانی $X$ و $Y$ از رابطهٔ (۸) محاسبه می‌شود:

$$COR(X,Y) = 1 - \frac{1}{d}\sum_{k=1}^{d}\rho\left(X^k, Y^k\right) \quad (8)$$

که در آن $\rho\left(X^k, Y^k\right)$ ضریب همبستگی پیرسون بین $k$-امین بعد $X$ و $Y$ است و d ابعاد هر سری است.

### ۳-۲-۵. معیار فاصله مبتنی‌بر شکل[5]

معیار فاصله مبتنی‌بر شکل (پاپاریزو و گراوانو[6]، ۲۰۱۷) در روش خوشه‌بندی k-shape استفاده می‌شود. در این روش براساس معیار همبستگی متقابل نرمال‌شده[7] معیار همبستگی متقابل با استفاده از الگوریتم تبدیل فوریه سریع[8] محاسبه می‌شود و با استفاده از میانگین هندسی خودهمبستگی هر سری زمانی نرمال‌سازی می‌شود. با توجه به دو بردار سری زمانی $X$ و $Y$ ، معیار فاصله مبتنی‌بر شکل با استفاده از رابطهٔ (۹) تعریف می‌شود.

1. Complexity Invariant Distance (CID)
2. Batista
3. Correlation-based distance (COR)
4. Chouakria & Nagabhushan
5. Shape-based Distance (SBD)
6. Paparrizos & Gravano
7. Normalized Cross-Correlation
8. Fast Fourier Transform

$$SBD(X,Y) = 1 - \max_{w}(NCC_w(X,Y)) \quad (9)$$

که در آن $w$ موقعیتی است که در آن $NCC_w(X,Y)$ حداکثر می‌شود (پاپاریزو و گراوانو، ۲۰۱۷).

در این مطالعه از یک روش مبتنی بر خوشه‌بندی ترکیبی برای شناسایی خوشه‌های مشتریان استفاده شده است که مبتنی‌بر مفاهیم گراف است (بونگون و یم-آن، ۲۰۱۸). دلیل استفاده از خوشه‌بندی ترکیبی وابستگی نتایج هر خوشه‌بندی به معیار فاصله انتخاب شده است. هرکدام از معیارهای فاصلهٔ سری زمانی از منظر خاصی عمل می‌کنند. برای این منظور استفاده از روش خوشه‌بندی ترکیبی در این پژوهش استفاده شد. الگوریتم ۱ (جدول ۲) فرایند خوشه‌بندی پیشنهادی را نشان می‌دهد.

همان‌طور که در الگوریتم ۱ مشخص است، فهرستی از الگوریتم‌های خوشه‌بندی سری زمانی روی داده‌های مشتریان اعمال می‌شوند. خروجی هر الگوریتم لیستی از برچسب‌های خوشهٔ هرکدام از مشتریان است. براساس لیست برچسب هر خوشه، ماتریس شباهت ساخته می‌شود. این ماتریس دارای ابعاد $n \times n$ بوده و هر درایهٔ $(i,j)$ آن نشان‌دهندهٔ آن است چند بار مشتری $i$ و مشتری $j$ در خوشهٔ یکسانی قرار گرفته‌اند. همان‌طور که در الگوریتم مشخص است این ماتریس به‌ازای خروجی هرکدام از الگوریتم‌های خوشه‌بندی مشارکت‌کننده، آپدیت می‌شود. درنهایت ماتریس شباهت نرمال‌سازی شده و هر درایه $(i, j)$ نمایانگر احتمال قرار گرفتن مشتری $i$ و مشتری $j$ در یک خوشه مشترک است. براساس ماتریس احتمال به‌دست‌آمده گرافی تشکیل می‌شود که رئوس آن نشان‌دهندهٔ یک سری زمانی و یالی از گره $i$ به گره $j$ دارای وزنی برابر با مقدار $(i,j)$ در ماتریس احتمال است. وظیفهٔ اصلی این رویکرد، حذف یال‌هایی است که مقادیر آن‌ها زیر آستانه[1] است که این امر به جدا شدن گره‌ها از هم منجر می‌شود. گره‌هایی که متصل باقی‌مانده‌اند، خوشه‌های نهایی را تشکیل می‌دهند.

**جدول ۲. الگوریتم خوشه بندی ترکیبی سری زمانی**

| Algorithm 1: Time Series Ensemble Clustering | |
|---|---|
| 1: | **Input:** Data: List of *N* Multivariate Time series of Customers |
| 2: | **Input:** Clustering_models: List of clustering models |
| 3: | **Input:** min_probability: Threshold for graph edge formation |
| 4: | **Input:** n_clusters: Number of Clusters |
| 5: | **Output:** Clustering_Results: Labels of the clustered data |
| 6: | similarity_matrix ← zeros(N,N) |
| 7: | **For each** model in clustering_models **do** |
| 8: | binary_matrix ← zeros(N,N) |
| 9: | cluster_labels ← model(data) |
| 10: | binary_matrix ← (Cityblock(cluster_labels, cluster_labels) == 0).astype(int) |
| 11: | similarity_matrix ← similarity_matrix + binary_matrix |
| 12: | **End For** |
| 13: | norm_sim_matrix ← similarity_matrix / similarity_matrix.diagonal() |
| 14: | graph ← (norm_sim_matrix > min_probability).astype(int) |
| 15: | Clustering_Results← connected_components(graph) |
| 16: | Return Clustering_Results |

## ۳-۳. تحلیل رفتاری مشتریان هر خوشه و برچسب‌گذاری به هر خوشه

برای تحلیل رفتاری مشتریان هر خوشه و برچسب‌گذاری به آن‌ها، از معیارهای مهمی مانند مجموع خرید و تکرار خرید که نشان‌دهندهٔ ارزش مشتری برای سازمان هستند، استفاده شده است. همچنین، روند مشتریان در سری‌های زمانی که می‌تواند صعودی، نزولی، یا پایدار باشد نیز در نظر گرفته می‌شود. این معیارها کمک می‌کنند تا مشتریان براساس میزان ارزش اقتصادی که برای شرکت ایجاد می‌کنند و الگوهای رفتاری آن‌ها در طول زمان، دسته‌بندی و برچسب‌گذاری شوند. درنتیجه، خوشه‌های مختلف مشتریان براساس تحلیل‌ها به سه دستهٔ اصلی تقسیم می‌شوند: مشتریان طلایی، نقره‌ای، و برنزی.

1. min_probability

### ۳-۴. استخراج ویژگی‌هایی ازسری‌های زمانی و فرموله کردن مسئلهٔ طبقه‌بندی با در نظر گرفتن برچسب خوشه به‌عنوان متغیر هدف و آموزش یک مدل مبتنی بر روش تقویت گرادیان

در این مرحله ویژگی‌هایی از سری زمانی استخراج شده و با در نظر گرفتن برچسب هر خوشه به‌عنوان برچسب مشتریان اعضای آن خوشه یک مدل طبقه‌بندی با روش تقویت گرادیان ساخته می‌شود. سپس، این مدل برای تحلیل اهمیت هر یک از ویژگی‌ها، به روش شاپلی منتقل می‌شود.

### ۴-۴.استفاده از روش شاپلی برای محاسبهٔ وزن ویژگی‌ها

در بسیاری از کاربردها، درک دلیل پیش‌بینی یک مدل می‌تواند به‌اندازهٔ دقت آن پیش‌بینی اهمیت داشته باشد. بااین‌حال، بیشترین دقت در مجموعه داده‌های بزرگ و مدرن بیشتر از مدل‌های پیچیده‌ای مانند مدل‌های یادگیری عمیق حاصل می‌شود که حتی کارشناسان نیز در تفسیر آن‌ها دچار مشکل می‌شوند. در پاسخ به این موضوع، روش توضیح افزودنی شاپلی (لوندبرگ، ۲۱۰۷) ارائه شده است که یک چارچوب یکپارچه برای تفسیر پیش‌بینی‌ها به شمار می‌آید.

شاپلی چارچوبی مبتنی‌بر نظریهٔ بازی‌ها برای تفسیر خروجی هر مدل یادگیری است که به محاسبهٔ سهم هر ویژگی در پیش‌بینی‌های مدل می‌پردازد. این روش براساس نظریهٔ بازی‌ها ساخته شده و از مقادیر شاپلی[1] استفاده می‌کند که از نظریهٔ بازی‌های مشارکتی آمده و به‌طور عادلانه امتیاز یک نتیجه را در میان یک مجموعه از مشارکت‌کنندگان (در اینجا ویژگی‌ها) توزیع می‌کند. ایدهٔ اصلی شاپلی محاسبهٔ مقادیر شاپلی برای هر ویژگی است. مقادیر شاپلی میانگین سهم یک ویژگی را در پیش‌بینی اندازه‌گیری می‌کند. این کار توزیع عادلانهٔ تأثیر هر ویژگی را تضمین می‌کند. به‌بیان‌دیگر، مقادیر شاپلی توضیح می‌دهند که هر ویژگی چگونه به حرکت پیش‌بینی مدل از میانگین (پایه) به خروجی واقعی برای یک نقطه داده خاص کمک می‌کند. شاپلی مستقل از مدل خاصی است و می‌تواند به هر مدل یادگیری ما شینی اعمال شود. یکی از مهمترین کاربردهای آن تعیین اهمیت ویژگی‌ها در یک مسئلهٔ پیش‌بینی یا طبقه‌بندی است.

### ۵-۴.در نظر گرفتن وزن ویژگی‌ها و استفاده از روش تاپسیس برای رتبه‌بندی مشتریان

در این مرحله رتبه‌بندی مشتریان در جایگاه یک مسئلهٔ تصمیم‌گیری چندمعیاره[2] (MCDM) با استفاده از ویژگی‌های کلیدی رفتاری مشتریان مدل‌سازی می‌شود. مسئلهٔ تصمیم‌گیری چندمعیاره در طیف گسترده‌ای از مسائل دنیای واقعی و برای قضاوت در مورد اهمیت ویژگی‌ها استفاده شده است (رهسپارفرد و مولایی[3]، ۲۰۱۹). برای انتخاب مجموعه بهینه مشتریان برای بهینه‌سازی استراتژی‌های بازاریابی، از روش تاپسیس استفاده شده است (هوانگ[4] و دیگران، ۱۹۸۱). این روش یکی از رویکردهای معتبر در زمینهٔ تصمیم‌گیری چندمعیاره است و برای انتخاب گزینه بهینه از میان چندین انتخاب به کار می‌رود. با توجه به‌سادگی و کارایی بالا، این روش مورد توجه بسیاری از محققان برای حل مسائل کلاسیک تصمیم‌گیری قرار گرفته است (هو[5] و دیگران، ۲۰۱۶). در مسائل تصمیم‌گیری، برخی از معیارها نشان‌دهندهٔ سود و برخی دیگر نشان‌دهندهٔ هزینه هستند و براساس اهمیت معیارها، به هرکدام از آن‌ها وزن اختصاص داده می‌شود. در روش تاپسیس، مفهوم راه‌حل ایده‌آل مثبت و منفی برای یافتن گزینهٔ بهینه به کار گرفته می‌شود. راه‌حل ایده‌آل مثبت بیشترین سود و کمترین هزینه را دارد، درحالی‌که راه‌حل ایده‌آل منفی بیشترین هزینه و کمترین سود را دارد. درنهایت، روش تاپسیس مشتریانی را انتخاب می‌کند که فاصلهٔ کم‌تری با راه‌حل ایده‌آل مثبت و فاصله بیشتری با راه‌حل ایده‌آل منفی دارند.

## ٤. مطالعهٔ موردی

در این بخش رویکردی پیشنهادی روی داده‌های یک شرکت فعال در صنایع غذایی اعمال شد و مشتریان این شرکت تحلیل شدند.

1. Shapley
2. Multi-Criteria Decision Making (MCDM)
3. RahseparFard & Molaei,
4. Hwang
5. Hu

### ۴-۱. مجموعه داده

در این مطالعه داده‌های فروش مربوط به یک شرکت فعال در حوزهٔ صنایع غذایی برای انجام آزمایش‌ها استفاده شده است که شامل تراکنش‌های مشتریان در طول پنج سال است. این مجموعه داده شامل اطلاعاتی مانند تاریخچهٔ خریدها، تعداد و نوع محصولات، ارزش و دفعات خرید است و ویژگی‌های مشتریان را نیز در برمی‌گیرد. تنوع محصولات و گستردگی اطلاعات این دیتاسِت امکان تحلیل دقیق‌تر رفتار مشتریان، شناسایی مشتریان باارزش، و بررسی روند خرید مشتریان در طول زمان را فراهم می‌کند. همچنین با استفاده از روش‌های پیشرفتهٔ تصمیم‌گیری چندمعیاره می‌توان اهمیت و رفتار مشتریان را تحلیل کرد.

### ۴-۲. پیش‌پردازش داده‌ها و تبدیل داده‌ها به سری زمانی چندمتغیره

داده‌های مربوط به رفتار خرید مشتریان به‌صورت خام جمع‌آوری شده بود. در ابتدا با در نظر گرفتن ابعاد مدل آر.اف.ام، داده‌ها به‌صورت ماهانه تجمیع شدند. از آنجا که تمام مشتریان انتخاب‌شده در همهٔ ماه‌ها خرید داشته‌اند، بنابراین مقدار ویژگی تازگی برای تمامی مشتریان صفر خواهد بود، به همین دلیل ویژگی R از فرایند تصمیم حذف شد. در نهایت برای هر مشتری سری زمانی با ابعاد ۵۶ برای ویژگی‌های M و F به دست آمد.

### ۴-۳. خوشه‌بندی سری‌های زمانی

#### ٤-۳-۱. انتخاب روش‌های خوشه‌بندی

همان‌طور که در بخش ۳ بیان شد، در این مطالعه از یک رویکرد خوشه‌بندی ترکیبی استفاده شده است. برای این منظور ابتدا براساس روش‌های خوشه‌بندی پایه، مشتریان به گروه‌های مشابه تقسیم می‌شوند. در این مطالعه، برای ایجاد تنوع در مدل‌های خوشه‌بندی، روش‌های خوشه‌بندی زیر استفاده شده‌اند:

۱- خوشه‌بندی k-means با معیار فاصله اقلیدسی

۲- خوشه‌بندی k-means با معیار DTW

۳- خوشه‌بندی k-shape

۴- خوشه‌بندی سلسله مراتبی با معیار CID

۵- خوشه‌بندی سلسله مراتبی با معیار COR

در تمامی الگوریتم‌های خوشه‌بندی تعداد خوشه‌ها برابر با ۳ در نظر گرفته شده است. دلیل این موضوع فرضیهٔ اولیه ما است که مشتریان می‌توانند روندهای رفتاری ثابت، صعودی، و نزولی داشته باشند. همچنین در الگوریتم خوشه‌بندی ترکیبی معیار حد آستانه برابر 0.8 در نظر گرفته شد.

#### ٤-۳-۲. سنجش کیفیت خوشه‌بندی

در این بخش کیفیت خوشه‌بندی ترکیبی با معیار شیلهوت[1] (روسوف[2]، ۱۹۸۷) بررسی شده است. به‌طور کلی معیار شیلهوت در بازهٔ [-1,+1] است و نزدیکی آن به عدد ۱ نشان‌دهندهٔ کیفیت بالای خوشه‌بندی است. این معیار برای روش خوشه‌بندی ترکیبی برابر با ۰٫۵۱ است که نشان می‌دهد خوشه‌بندی ترکیبی توانسته است مشتریان با رفتارهای مشابه را شناسایی کند. همچنین در یک مطالعهٔ خوشه‌بندی اندازهٔ خوشه‌ها از جنبهٔ کاربردهای دنیای واقعی اهمیت زیادی دارد. در جدول ۳ اندازهٔ خوشه‌ها برای الگوریتم‌های خوشه‌بندی مختلف نمایش داده شده است. همان‌طور که در جدول آمده است، روش پیشنهادی داده‌ها را در ۳ خوشه به‌طور متوازن توزیع می‌کند. این موضوع از لحاظ کاربرد عملی از اهمیت ویژه‌ای برخوردار است چرا که اندازهٔ خوشه‌ها در یک برنامه بازاریابی باید تا حدودی معقولانه باشد و مشتریان در خوشه‌های مختلف توزیع شوند. همچنین نتایج جدول ۳ نشان‌دهندهٔ این موضوع است که برخی روش‌های خوشه‌بندی منفرد در انجام خوشه‌بندی ضعیف عمل کرده‌اند و نتوانسته‌اند داده‌ها را به‌خوبی افراز کنند.

1. Silhouette
2. Rousseeuw

**جدول ۳. سنجش کیفیت خوشه‌بندی با استفاده از توزیع داده‌ها در خوشه‌ها**

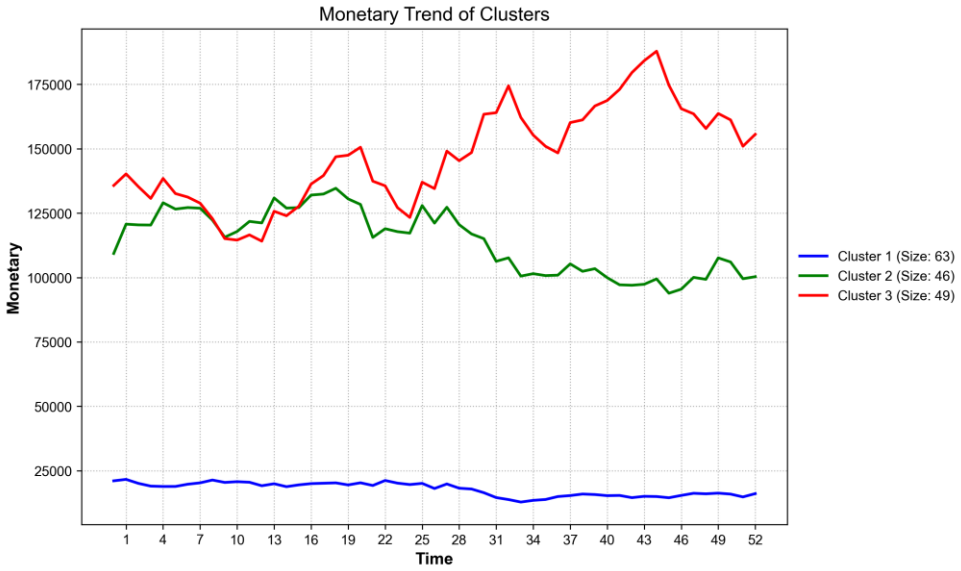
| روش خوشه‌بندی | اندازه خوشه ۱ | اندازه خوشه ۲ | اندازه خوشه ۳ |
|---|---|---|---|
| خوشه‌بندی k-means با معیار فاصله اقلیدسی | ۱۵۳ | ۳ | ۲ |
| خوشه‌بندی k-means با معیار DTW | ۱۲۹ | ۲۴ | ۵ |
| خوشه‌بندی k-shape | ۱۰۸ | ۴۹ | ۱ |
| خوشه‌بندی سلسله مراتبی با معیار CID | ۷۳ | ۷۰ | ۱۵ |
| خوشه‌بندی سلسله مراتبی با معیار COR | ۱۵۴ | ۲ | ۲ |
| روش پیشنهادی | ۶۳ | ۴۶ | ۴۹ |

## ۴-۴. تحلیل خوشه‌های به‌دست‌آمده

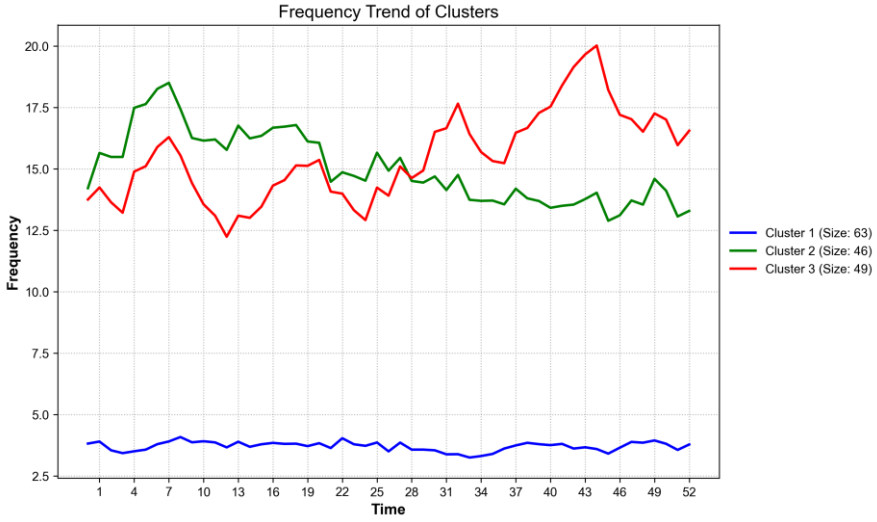
با اعمال الگوریتم خوشه‌بندی ترکیبی، خوشه‌های مشتریان به دست آمده و سپس میانگین خوشه‌ها ترسیم شد. برای تحلیل مجزای ویژگی‌های M و F، سری میانگین هرکدام از این متغیرها در نمودارهای جداگانه به‌ترتیب در شکل‌های ۳ و ۴ نمایش داده شده است. همان‌طور که در شکل ۳ مشخص است مشتریان خوشهٔ ۱ به لحاظ مقدار M رفتار تقریباً ثابت و پایداری را از خود نشان داده‌اند. این مشتریان نسبت به مشتریان دو خوشهٔ دیگر مبلغ خرید کمتری داشته‌اند. همچنین در ماه‌های آخر، مبلغ خرید آن‌ها کاهش اندکی داشته است. این موضوع برای متغیر F این مشتریان نیز صادق است. همان‌طور در شکل ۴ نشان داده شده است روند رفتاری این مشتریان از منظر F روند تقریباً ثابتی بوده است. این مشتریان را می‌توان جزء خوشهٔ برنزی نظر گرفت. ثبات در روند این خوشه می‌تواند به‌معنای آن باشد که این مشتریان در طول زمان به‌طور پایدار خرید خود را حفظ کرده‌اند، اما ارزش خرید پایین‌تری نسبت به سایر خوشه‌ها دارند.

مشتریان خوشهٔ ۲ در نمودار شکل ۳، تا ماه ۷ روند نسبتاً صعودی را به لحاظ M داشته‌اند، ولی در بازهٔ ماه هفتم تا سیزدهم روند نزولی برای این خوشه مشاهده می‌شود. بررسی داده‌های خرید بیانگر آن است که این بازهٔ مربوط به همه‌گیری ویروس کرونا در ایران بوده است که تأثیر خود را در مشتریان خوشهٔ ۲ به این صورت نشان داده است. همچنین تحلیل روند رفتاری از منظر متغیر برای مشتریان خوشهٔ ۲ تائیدکنندهٔ این موضوع است. از ماه سیزدهم تا ماه هجدهم یک روند صعودی در رفتار مشتریان خوشهٔ ۲ در متغیر M و F دیده می‌شود. از ماه هجدهم به بعد تقریباً یک روند نزولی از منظر هر دو متغیر M و F برای خوشهٔ ۲ دیده می‌شود که در مورد علت آن بایستی تحلیل بیشتری انجام شود. مشتریان این خوشه را می‌توان جزء خوشهٔ نقره‌ای در نظر گرفت. به‌طورکلی می‌توان گفت مشتریان خوشهٔ ۲ از مشتریان با برچسب طلایی تبدیل به مشتریان با برچسب نقره‌ای شده‌اند؛ و در صورتی‌که که اقدامات بازاریابی خاص این خوشه اتفاق انجام نشود احتمال ریزش این مشتریان وجود دارد.

همان‌طور که در شکل ۳ مشاهده می‌شود، مشتریان خوشهٔ ۳ در ابتدا روند صعودی داشته‌اند، ولی مشابه مشتریان خوشهٔ ۲ از ماه هفتم تا سیزدهم روند نزولی داشته‌اند که به نظر می‌رسد اتفاقات کووید-۱۹ روی رفتار خرید این مشتریان نیز تأثیرگذار بوده است. به‌طورکلی این خوشه در مقادیر M طی زمان روند افزایشی دارد. باوجود نوسانات، به‌طورکلی یک افزایش در طول زمان مشاهده می‌شود. مقادیر M در این خوشه در بازهٔ ۱۰۰٬۰۰۰ تا ۱۷۵٬۰۰۰ متغیر است. چندین قله و دره در منحنی روند این خوشه وجود دارد که نشان‌دهندهٔ دوره‌های ناپایداری است. در حدود هفتهٔ ۳۷، مقدار M به قله‌ای در حدود ۱۷۵٬۰۰۰ می‌رسد و پس از آن یک کاهش جزئی رخ می‌دهد و سپس دوباره افزایش می‌یابد. علی‌رغم نوسانات، خوشهٔ ۳ میانگین M بالاتری نسبت به خوشه‌های ۱ و ۲ دارد. همچنین مطابق نمودار F مربوط به این خوشه، از منظر F نیز این خوشه رفتاری مشابه متغیر M از خود نشان می‌دهد و به‌طورکلی روند افزایشی دارد. بنابراین اعضای این خوشه را می‌توانیم جز مشتریان طلایی لحاظ کنیم.




**شکل ۳. نمودار روند خوشه‌های مشتریان براساس متغیر میزان خرید مشتریان در طول زمان**

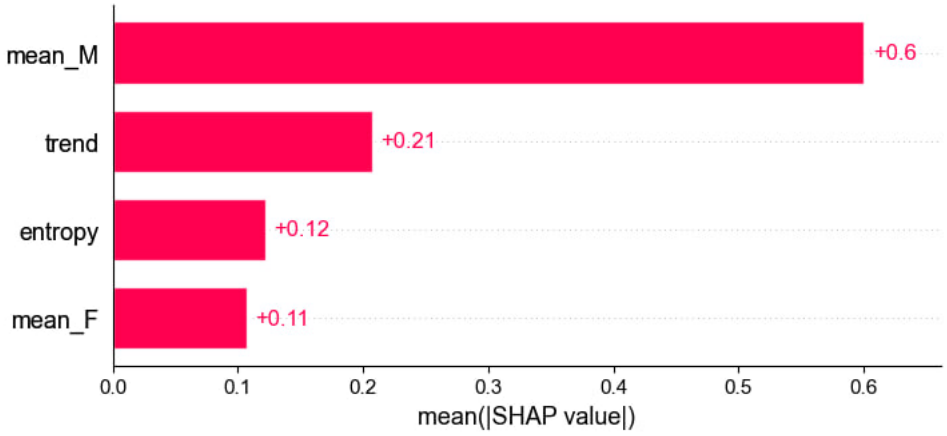



**شکل ۴. نمودار روند خوشه‌های مشتریان براساس متغیر تکرار خرید مشتریان در طول زمان**

## ۴-۵. ساخت مدل طبقه‌بندی و استفاده از روش شاپلی برای استخراج اهمیت هر ویژگی

با توجه به این‌که خوشهٔ ایده‌آل ما، خوشهٔ شماره ۳ است، میانگین قدرِ مطلق مقادیر شاپلی برای هر یک از ویژگی‌ها با تحلیل شاپلی روی هر یک از نمونه‌های موجود در خوشهٔ ۳ محاسبه می‌شود. با در نظر گرفتن برچسب هر خوشه برای تمامی اعضای آن خوشه، در اینجا یک مدل طبقه‌بندی با استفاده از روش تقویت گرادیان ساخته می‌شود. ویژگی‌های ورودی شامل ویژگی آنتروپی، میانگین M، میانگین F و مؤلفهٔ روند هر سری است که با استفاده از کتابخانه tsfeature استخراج می‌شوند. مدل تقویت گرادیان این ویژگی‌ها را به متغیر هدف (برچسب خوشه) نگاشت می‌کند. با ساخت مدل دسته‌بندی می‌توانیم در مورد دلایل مربوط به تعلّق هر مشتری به هر خوشه اطلاعات بیشتری کسب کنیم و به تحلیل ویژگی‌ها و رفتارهای آن‌ها در هر خوشه بپردازیم. برای این منظور، از روش شاپلی برای محاسبهٔ مقادیر شاپلی هر ویژگی استفاده می‌کنیم. مقادیر شاپلی میانگین سهم یک ویژگی را در پیش‌بینی اندازه‌گیری کرده و ترکیب‌های ممکن ویژگی‌ها را در نظر می‌گیرد.

شکل ۵، میانگین قدرِ مطلق مقادیر شاپلی هر یک از متغیرها را نشان می‌دهد. همان‌طور که در شکل مشخص است، ویژگی میانگین M، بیشترین تأثیر را در طبقه‌بندی نمونه‌ها در خوشهٔ ۳ داشته است. مقادیر شاپلی به‌دست‌آمده نرمال‌سازی شده و درنهایت، مقادیر وزن هر ویژگی برای استفاده در روش تاپسیس محاسبه می‌شود.




**شکل ۵. میانگین قدرِ مطلق مقادیر شاپلی برای ویژگی‌های کلیدی مشتریان**

## ۴-۶. به‌کارگیری تاپسیس

در این بخش، مطابق روش تاپسیس ابتدا ماتریس تصمیم ($D$) که شامل n سطر و m ستون تشکیل می‌شود، به‌طوری‌که هر سطر نمایانگر یک مشتری و هر ستون نشان‌دهندهٔ یکی از ویژگی‌های مثبت یا منفی است. در این پژوهش چهار ویژگی معیار میانگین ارزش خرید (Mean_m)، میانگین تکرار خرید (Mean_F)، روند(Trend) ، و آنتروپی (Entropy) برای اولویت‌بندی مشتریان انتخاب شدند. معیارهای میانگین ارزش خرید، تکرار خرید، و روند به‌عنوان معیارهای مثبت و آنتروپی به‌عنوان معیار منفی در نظر گرفته شدند. پس از نرمال‌سازی مقادیر مربوط به این معیارها، راه‌حل ایده‌آل مثبت ($A^+$) و راه‌حل ایده‌آل منفی ($A^-$) تعیین شد. سپس، فاصلهٔ هر مشتری از ایده‌آل مثبت ($S^+$) و ایده‌آل منفی ($S^-$) محاسبه و در نهایت، امتیاز نهایی هر مشتری ($C^*$) برای رتبه‌بندی محاسبه شد. مشتریانی که فاصلهٔ کمتری از ایده‌آل مثبت و فاصلهٔ بیشتری از ایده‌آل منفی داشتند، امتیاز بالاتری کسب کردند. نمونه‌ای از نتایج حاصل از اجرای روش تاپسیس برای ده مشتری برتر در جدول ۴ گزارش شده است.

**جدول ۴. نمونه‌ای از نتایج تاپسیس برای رتبه‌بندی مشتریان**

| Customer_ID | Mean_M | Mean_F | Trend | Entropy | $S^+$ | $S^-$ | $C^*$ |
|---|---|---|---|---|---|---|---|
| 72 | 0.39501 | 0.32585 | 0.07621 | 0.07350 | 0.02573 | 0.23866 | 0.90269 |
| 26 | 0.36544 | 0.32477 | 0.03278 | 0.08602 | 0.03825 | 0.22068 | 0.85229 |
| 78 | 0.32862 | 0.26684 | 0.09985 | 0.06173 | 0.04776 | 0.19890 | 0.80636 |
| 4 | 0.27860 | 0.45465 | 0.13676 | 0.05826 | 0.07038 | 0.17527 | 0.71350 |
| 22 | 0.26324 | 0.23974 | 0.11426 | 0.06430 | 0.08361 | 0.16037 | 0.65732 |
| 48 | 0.18333 | 0.16189 | 0.02888 | 0.08466 | 0.13476 | 0.11001 | 0.44945 |
| 23 | 0.16894 | 0.15021 | 0.11858 | 0.06791 | 0.14028 | 0.10416 | 0.42611 |
| 51 | 0.15731 | 0.13938 | 0.09782 | 0.06313 | 0.14774 | 0.09630 | 0.39460 |
| 33 | 0.15895 | 0.14680 | 0.02290 | 0.05928 | 0.14920 | 0.09532 | 0.38983 |
| 32 | 0.15553 | 0.14321 | 0.05834 | 0.07805 | 0.14986 | 0.09386 | 0.38512 |

با توجه به جدول ۴ می‌توان تحلیل‌های مهمی را دربارهٔشناسایی جامعهٔ هدف استراتژی بازاریابی و بررسی معیارهای مختلف مشتریان ارائه داد. در این نمونه، براساس امتیاز نهایی ($C^*$)، مشتریان با شناسهٔ ۷۲، ۲۶، و ۷۸، به‌عنوان بهترین گزینه‌ها برای هدف‌گذاری در ا ستراتژی بازاریابی شنا سایی می شوند. این م شتریان به‌ویژه در معیارهای میانگین ارزش خرید و تکرار خرید عملکرد

خوبی داشته و نزدیک‌ترین فاصله به ایده‌آل مثبت ($S^+$) و دورترین فاصله از ایده‌آل منفی ($S^-$) را دارند. بنابراین، تمرکز روی این مشتریان می‌تواند به افزایش درآمد و حفظ مشتریان وفادار منجر شود.

مشتری ۷۲ با امتیاز ۰٫۹۰۲۶۹، بهترین عملکرد را در بین مشتریان دارد و در تمامی معیارها ازجمله میانگین ارزش خرید، تکرار خرید، و به‌ویژه در معیار روند در مقایسه با سایرین شرایط بهتری دارد. این مشتری به‌طور مکرر خرید می‌کند و در حال حاضر نیز روند صعودی دارد که می‌تواند به‌عنوان نشانه‌ای از پتانسیل بالای او برای خریدهای آینده در نظر گرفته شود. درحالی‌که مشتری ۳۲ که امتیاز ۰٫۳۸۵۱۲ را کسب کرده، نیاز به توجه بیشتری دارد. این نوع از مشتریان ممکن است به دلیل خریدهای کمتر یا ارزش خرید پایین، در معرض خطر ریزش قرار گیرند. ازاین‌رو، طراحی برنامه‌های خاص برای جذب دوبارهٔ این دسته از مشتریان می‌تواند به سازمان کمک کند. درنهایت، نتایج تحلیل‌ها روی این ده نمونه مشتری نشان می‌دهد که مشتریان ۷۲، ۲۶، و ۷۸ می‌توانند به‌عنوان جامعهٔ هدف اصلی برای استراتژی بازاریابی انتخاب شوند و مشتریان با امتیاز پایین‌تر باید زیر نظر قرار گیرند تا از خطر ریزش آن‌ها پیشگیری شود.

نتایج این پژوهش با پژوهش پیشین در خوشه‌بندی و تحلیل رفتار خرید مشتریان مطابقت دارد و گامی برای توسعهٔ رویکردهای پیشین است. پژوهش‌های قبلی بر اهمیت استفاده از روش آر.اف.ام در تقسیم‌بندی مشتریان تأکید کرده‌اند و بررسی کرده‌اند که روند میزان خرید و تکرار خرید مشتریان فاکتوری مهم در بررسی رفتار مشتریان هستند (عباسی مهر و بحرینی، ۲۰۲۲؛ اسمعیلی و حکیمی، ۲۰۲۳) همچنین جایگاه مهم سری‌های زمانی برای بررسی رفتار مشتریان در طول زمان در مطالعات قبلی بررسی شده است (عباسی مهر و شبانی ۲۰۲۱ ب). همچنین روش‌های تصمیم‌گیری چند معیاره مانند تاپسیس برای اولویت‌بندی مشتریان هدف سازمان‌ها استفاده شده است (هو و همکاران،۲۰۱۶). در مطالعهٔ حاضر نیز، علاوه‌بر استفاده از روش‌های یاد شده، از سری‌های زمانی چندمتغیره برای تحلیل روند رفتار مشتریان استفاده شده است که در مطالعات پیشین کمتر به آن‌ها توجه شده است.. علاوه‌بر این، ساخت مدل طبقه‌بندی، استفاده از روش شاپلی برای استخراج اهمیت هر ویژگی و اولویت‌بندی مشتریان با روش تاپسیس باعث شده تحلیل بهتر و واقعی از رفتار خوشه‌های مختلف مشتریان ارائه شود. درنتیجه، روش ارائه‌شده در این مقاله با گسترش تئوری‌های موجود در ادبیات پژوهش، رویکردی جدید ارائه داده است که به بهبود تحلیل رفتار مشتری و طراحی استراتژی‌های بازاریابی کمک می‌کند.

## ۵. نتیجه‌گیری

در این پژوهش، رویکرد جدیدی برای تحلیل رفتار پویای مشتریان از طریق ترکیب مدل آر.اف.ام و خوشه‌بندی سری‌های زمانی چندمتغیره ارائه شده است که می‌تواند به بهبود طراحی استراتژی بازاریابی و مدیریت منابع سازمانی کمک کند. براساس تحلیل نتایج، می‌توان مشتریان شرکت فعال در حوزهٔ صنایع غذایی را به سه دسته اصلی تقسیم کرد: مشتریان با سطح خرید بالا و روند صعودی (خوشه طلایی)، مشتریان با سطح خرید بالا اما روند نزولی (خوشه نقره‌ای)، و مشتریان با سطح خرید پایین و روند پایدار (خوشه برنزی).

برای درک بهتر متغیرهای مؤثر بر رفتار مشتریان و اولویت‌بندی آن‌ها، بررسی جامعی انجام شد. نتایج اولویت‌بندی نشان می‌دهد که مشتریانی که ازنظر رفتار خرید در رتبه‌های بالاتری قرار دارند، پتانسیل بیشتری برای ایجاد درآمد و افزایش سودآوری سازمان دارند. به‌ویژه، مشتریانی که دارای خریدهای مکرر و ارزش خرید بالا با روند صعودی هستند (گروه طلایی)، می‌توانند به‌عنوان اهداف اصلی استراتژی‌های بازاریابی مؤثر انتخاب شوند. این یافته‌ها، سازمان‌ها را قادر می‌سازد تا با تحلیل رفتار مشتریان، تصمیم‌گیری‌های درست‌تری داشته باشند و مزیت رقابتی خود را در بازار افزایش دهند.

روش ارائه شده در این مقاله قابلیت استفاده در مجموعه داده‌های مختلف را دارد و می‌تواند رفتار خرید مشتریان را در صنایع و شرایط مختلفی نشان دهد. توسعهٔ این روش برای پیش‌بینی الگوهای خرید آینده مشتریان با در نظر گرفتن متغیرهای بیشتر می‌تواند زمینهٔ پژوهش آتی باشد. همچنین ترکیب روش پیشنهادی با ابزارهای هوش مصنوعی و یادگیری عمیق می‌تواند دقت و جامعیت نتایج حاصل از تحلیل رفتار خرید مشتریان را بهبود بخشد.

## ٦. فهرست منابع


رهسپارفرد، خ.، و مولایی، ر. (۱۳۹۷). بررسی چالش‌های اینترنت اشیا با استفاده از روش مدل‌سازی ساختاری تفسیری. علوم و فنون مدیریت اطلاعات.۴ (۴)،۸۲-۶۳

شریفی، ح.، شهسواری پور، ن.، شریفی فینی، م.، وبهمن یار، ر. (۱۴۰۲). ارائه رویکردی مبتنی بر سابقه خرید مشتریان و توصیه‌ی محصول به مشتری: مورد مطالعه مشتریان دیجی‌کالا. چشم‌انداز مدیریت صنعتی, ۱۳(۲)،۹۹-۱۱۸.

قربانیان، ع.، و رضوی، ح. (۱۴۰۳). یک رویکرد جدید به‌منظور خوشه‌بندی سری‌های زمانی با استفاده از ترکیب زیرسری‌هایی زمانی. مهندسی صنایع و مدیریت. ۴۰ (۱)،۴۱-۲۷

ولایتی، م.، حسین زاده لطفی، ف.، شهریاری، م.، و رهنمای رود پشتی، ف. (۱۳۹۷). رویکرد داده‌کاوی در بخش‌بندی بازار مشتریان به منظور اتخاذ استراتژی‌های کارا (مطالعه موردی صنعت مخابرات). اقتصاد مالی، ۱۱(۴۱)، ۲۴۳-۲۶۶.

یوسفی زاد، ا.، و ثریایی، ع. (۱۳۹۸). بررسی و خوشه‌بندی مشتریان، بر اساس مدل RFM و طراحی الگویی برای ارائه خدمات به مشتریان کلیدی. پژوهشنامه مدیریت اجرایی، ۱۰(۲۰)، ۱۷۵-۱۹.

Abbasimehr, H., & Baghery, F. S. (2022). A novel time series clustering method with fine-tuned support vector regression for customer behavior analysis. *Expert Systems with Applications, 204*, 117584.

Abbasimehr, H., & Bahrini, A. (2022). An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. *Expert Systems with Applications, 192*, 116373. doi:https://doi.org/10.1016/j.eswa.2021.116373.

Abbasimehr, H., & Shabani, M. (2019). *Forecasting of customer behavior using time series analysis.* Paper presented at the The 7th International Conference on Contemporary Issues in Data Science.

Abbasimehr, H., & Shabani, M. (2021a). A new framework for predicting customer behavior in terms of RFM by considering the temporal aspect based on time series techniques. *Journal of Ambient Intelligence and Humanized Computing, 12*(1), 515-531. doi:10.1007/s12652-020-02015-w

Abbasimehr, H., & Shabani, M. (2021b). A new methodology for customer behavior analysis using time series clustering: A case study on a bank's customers. *Kybernetes*, 50(2), 221-242.

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering–a decade review. *Information systems*, 53, 16-38.

Akhondzadeh-Noughabi, E., & Albadvi ,A. (2015). Mining the dominant patterns of customer shifts between segments by using top-k and distinguishing sequential rules. *Management Decision*, 53(9), 1976-2003.

Al-Naymat, G., Chawla, S., & Taheri, J. (2012). Sparsedtw: A novel approach to speed up dynamic time warping. *arXiv preprint arXiv:1201.2969*.

Alqurashi, T., & Wang, W. (2019). Clustering ensemble method. *International Journal of Machine Learning and Cybernetics*, 10, 1227-1246.

Alsayat, A. (2023). Customer decision-making analysis based on big social data using machine learning: a case study of hotels in Mecca. *Neural Computing and Applications*, 35(6), 4701-4722.

Barough, S. S., Safavi-Naini, S. A. A., Siavoshi, F., Tamimi, A., Ilkhani, S., Akbari, S., . . . Pourhoseingholi, M. A. (2023 .(Generalizable machine learning approach for COVID-19 mortality risk prediction using on-admission clinical and laboratory features. *Scientific Reports, 13*(1), 2399.

Batista, G. E., Keogh, E. J., Tataw, O. M., & De Souza, V. M. (2014). CID: an efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery, 28*, 634-669.

Boongoen, T., & Iam-On, N. (2018). Cluster ensembles: A survey of approaches with recent extensions and applications. *Computer Science Review, 28*, 1-25.

Chouakria, A. D., & Nagabhushan, P. N. (2007). Adaptive dissimilarity index for measuring time series proximity. *Advances in Data Analysis and Classification, 1*, 5-21.

Coussement, K., Van den Bossche, F. A., & De Bock, K. W. (2014). Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees. *Journal of Business Research, 67*(1), 2751-2758.

Dhanushkodi, K., Bala, A., Kodipyaka, N., & Shreyas, V. (2024). Customer Behaviour Analysis and Predictive Modelling in Supermarket Retail: A Comprehensive Data Mining Approach. *IEEE Access*.

Doğan, O., Ayçin, E., & Bulut, Z. (2018). Customer segmentation by using RFM model and clustering methods: a case study in retail industry. *International Journal of Contemporary Economics and Administrative Sciences, 8*.

Duda, R., Hart, P., Stork, D., & Ionescu, A. (2000). *Pattern classification, chapter nonparametric techniques*. Wiley-Interscience Publication.

Ebadi Jalal, M., & Elmaghraby, A. (2024). Analyzing the Dynamics of Customer Behavior: A New Perspective on Personalized Marketing through Counterfactual Analysis. *Journal of Theoretical and Applied Electronic Commerce Research, 19*(3), 1660-1681.

Ghorbanian, A., & Razavi, H., ideh. (2024). A NEW APPROACH TO TIME SERIES CLUSTERING BY COMBINATION OF SUB-SERIES. Sharif Journal of Industrial Engineering & Management, 40(1), 27-41. doi:10.24200/j65.2022.60405.2303. [In Persian]

Guerola-Navarro, V., Oltra-Badenes, R., Gil-Gomez, H., & Gil-Gomez, J-.A. (2020). Customer relationship management (CRM): a bibliometric analysis. *International Journal of Services Operations and Informatics, 10*(3), 242-268.

Heldt, R., Silveira, C. S., & Luce, F. B. (2021). Predicting customer value per product: From RFM to RFM/P. *Journal of Business Research, 127*, 444-453.

Hu, J., Du, Y., Mo, H., Wei, D., & Deng, Y. (2016). A modified weighted TOPSIS to identify influential nodes in complex networks. *Physica A: Statistical Mechanics and its Applications, 444*, 73-85.

Hu ,Y.-H., Huang, T. C.-K., & Kao, Y.-H. (2013). Knowledge discovery of weighted RFM sequential patterns from customer sequence databases. *Journal of Systems and Software, 86*(3), 779-788. doi:https://doi.org/10.1016/j.jss.2012.11.016

Hughes, A. M. (2005). *Strategic database marketing*.McGraw-Hill Pub. Co.

Hwang, C.-L., Yoon, K., Hwang, C.-L., & Yoon, K. (1981). Methods for multiple attribute decision making. *Multiple attribute decision making: methods and applications a state-of-the-art survey*, 58-191.

Hyndman, R. J., Wang, E., & Laptev, N. (2015). *Large-scale unusual time series detection.* Paper presented at the 2015 IEEE international conference on data mining workshop (ICDMW).

Iam-On, N., Boongoen, T., & Garrett, S. (2010). LCE: a link-based cluster ensemble method for improved gene expression data analysis. *Bioinformatics, 26*(12), 1513-1519.

KATRAGADDA, V. (2022). Dynamic Customer Segmentation: Using Machine Learning to Identify and Address Diverse Customer Needs in Real-Time. *IRE Journals, 5*(10), 278 .۲۷۹-

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.

Kim, E.-Y., Hwang, D.-U., & Ko, T.-W. (2012). Multiscale ensemble clustering for finding modules in complex networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 85*(2), 026119.

Liu, Y., & Chen, C. (2022). Improved RFM model for customer segmentation using hybrid meta-heuristic algorithm in medical IoT applications. *International Journal on Artificial Intelligence Tools, 31*(01), 2250009.

Lundberg, S. (2017). A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*.

Montero, P., & Vilar, J. A. (2015). TSclust: An R package for time series clustering. *Journal of Statistical Software, 62*, 1-43.

Mosaddegh, A., Albadvi, A., Sepehri, M. M., & Teimourpour, B. (2021). Dynamics of customer segments: A predictor of customer lifetime value. *Expert Systems with Applications, 172*, 114606. doi:https://doi.org/10.1016/j.eswa.2021.114606

Pakzad, S. S., Roshan, N., & Ghalehnovi, M. (2023). Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. *Scientific Reports, 13*(1), 3646.

Paparrizos, J., & Gravano, L. (2017). Fast and accurate time-series clustering. *ACM Transactions on Database Systems (TODS), 42*(2), 1-49.

Parvaneh, A., Abbasimehr, H., & Tarokh, M. J. (2012). Integrating AHP and data mining for effective retailer segmentation based on retailer lifetime value. *Journal of Optimization in Industrial Engineering, 5*(11), 25-31.

Peter, M., Mofi, H., Likoko, S., Sabas, J., Mbura, R., & Mduma, N. (2025). Predicting customer subscription in bank telemarketing campaigns using ensemble learning models. *Machine Learning with Applications*, 100618.

Priyambada, S. A., Er, M., Yahya, B. N., & Usagawa, T. (2021). Profile-based cluster evolution analysis: Identification of migration patterns for understanding student learning behavior. *IEEE Access, 9*, 101718-101728.

RahseparFard, K., & Molaei, R. (2019). investigating Challenges on the Internet of things by Using Interpretive Structural Modeling. *Sciences and Techniques of Information Management, 4*(4), 63-82.[In Persian]

Ramasso, E., Placet, V., & Boubakar, M. L. (2015). Unsupervised consensus clustering of acoustic emission time-series for robust damage sequence estimation in composites. *IEEE Transactions on Instrumentation and Measurement, 64*(12), 3297-3307.

Ramezani, F., Parvez, S., Fix, J. P., Battaglin, A., Whyte, S., Borys, N. J., & Whitaker, B. M. (2023). Automatic detection of multilayer hexagonal boron nitride in optical images using deep learning-based computer vision. *Scientific Reports, 13*(1), 1595.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics, 20*, 53-65.

Sharifi Esfahani, H., Shahsavari Pour, N., Sharifi Fini, M. H., & Bahmanyar, R. (2023). A Customer-Centric Approach for Recommending Products: A Case Study of Digikala. Journal of Industrial Management Perspective, 13(2), 99-118. doi:10.48308/jimp.13.2.99 . [In Persinan]

Smaili, M. Y., & Hachimi, H. (2023). New RFM-D classification model for improving customer analysis and response prediction. *Ain Shams Engineering Journal, 14*(12), 102254.

Song, M., Zhao, X., E, H., & Ou, Z. (2016). *Statistic-based CRM approach via time series segmenting RFM on large scale data.* Paper presented at the Proceedings of the 9th International Conference on Utility and Cloud Computing.

Sun, Y., Liu, H., & Gao, Y. (2023). Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model. *Heliyon, 9*(2).

Uosefy Zad, A., & Sorayaei, A. (2019). Inspecting the Effective Factors on Identification and Maintenance of Key Customers Based on RFM Model and Designing a Model for Providing Services. *Journal of Executive Management, 10*(20), 175-198. doi:10.22080/jem.2019.15814.2834. [In Persian]

Velayati, M., HosseinzadehLotfi, F., Shahriari, M., & rahnamaroodposhti, F. (2018). A data mining approach for customer segmentation for making effective strategies ( Case study of telecom). *Financial Economics,11*(41), 243-266. [In Persian]

Wang, S., Sun ,L., & Yu, Y. (2024). A dynamic customer segmentation approach by combining LRFMS and multivariate time series clustering. *Scientific Reports, 14*(1), 17491.

Wei, J. T., Lin, S.-Y., Yang, Y.-Z., & Wu, H.-H. (2019). The application of data mining and RFM model in market segmentation of a veterinary hospital. *Journal of Statistics and Management Systems, 22*(6), 1049-1065.

Yang, Y., & Jiang, J. (2014). HMM-based hybrid meta-clustering ensemble for temporal data. *Knowledge-Based Systems, 56*, 299-310.