مجله تحقیق در عملیات در کاربردهای آن، سال بیست و دوم، شماره اول (پیاپی ۸۴)،بهار ۰۴، ص ص ۲۳-۱
شاپا ۷۲۸۶-۲۲۵۱
مقاله مروری

پیش‌بینی الگوی پایداری مشتریان در صنعت خودرویی و شرکت‌های خودروساز با استفاده از سیستم خبره

اعظم السادات نوربخش[1*]، سیدمیلاد میرمحمدیان[2]، محمدحسین مهدی زاده نادری[3]
۱-استادیار، گروه کامپیوتر و فناوری اطلاعات، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران
۲-کارشناس ارشد، کارشناس شبکه مرکز آمار و فناوری اطلاعات، قوه قضاییه، تهران، ایران
۳- کارشناس ارشد، مدیر امنیت شبکه مرکز آمار و فناوری اطلاعات، قوه قضاییه، تهران، ایران

رسید مقاله: ۲۹ اسفند ۱۴۰۱
پذیرش مقاله: ۲۳ مرداد ۱۴۰۲

چکیده
در دهه‌های اخیر، با توجه افزایش هزینه‌های جذب مشتری جدید به طور پیوسته، توجه به نگهداری مشتریان و بالابردن وفاداری آن‌ها، برای سودآوری سازمان‌ها بسیار مهم و حساس است. از این رو سازمان‌ها برنامه‌های مختلفی را برای افزایش ماندگاری مشتریان با ارزش خود (مشتریانی با اتلاف منبع کمتر و سودآوری بالا) اجرا می‌کنند. پژوهش حاضر با درنظر گرفتن قابلیت‌های داده کاوی در مدیریت و طراحی به پیاده‌سازی یک مدل پیش‌بینی رفتار رویگردانی مشتریان در حوزه صنعت، با بهره‌گیری از روش‌شناسی استاندارد CRISP-DM بر اساس مدل RFM و تکنیک‌های جنگل تصادفی و درخت‌های فزاینده، به جستجو در پایگاه داده مشتریان یک شرکت خودروساز، که بیش از یک قرارداد خرید محصول با آن شرکت خودروساز داشته‌اند، پرداخته است. با کاربرد مدلی بر مبنای تکنیک‌های جنگل تصادفی،درخت‌های فزاینده و مدل پیش‌بینی ترکیبی، مشتریانی که تمایل به رویگردانی دارند شناسایی‌شده و راهکارهای بازاریابی موثر برای این گروه برنامه‌ریزی می‌شود. تحلیل رفتار مشتری نشان می‌دهد که طول ارتباط فعال مشتری، تناوب خرید نسبی و متوسط فاصله زمانی بین خرید از بهترین پیش‌بینی‌کنندگان می‌باشند. همچنین تکنیک پیش بینی ترکیبی نسبت به تکنیک‌های جنگل تصادفی و درخت‌های فزاینده نتایج بهتری را نشان داده است.

کلمات کلیدی: رویگردانی مشتریان، داده کاوی، مدل RFM ، صنعت خودرو.

۱ مقدمه
رشد صنعت خودرو به عنوان یکی از شاخص‌های توسعه یافتگی از سویی به عنوان یکی از عمده‌ترین نهادها در بازار مالی هر اقتصادی مطرح بوده و از دیگر سوی فعالیت دیگر نهادها را پشتیبانی می‌کند. نقش صنعت خودرو

* عهده‌دار مکاتبات
آدرس الکترونیکی: Nourbakhsh@iau.ac.ir

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]
[DOI: 10.71773/jamlu-2025-1-2127]

ایجاب می‌کند که کارا و اثربخش بودن آن مورد توجه خاص سیاست‌گذاران قرار گیرد. بدین سبب توجه به کارکرد صحیح بازار خودرو و تاکید بر رقابتی بودن و عوامل موثر بر رقابت‌پذیری آن حایز اهمیت است. از ویژگی‌های خاص این صنعت این است که در بازار خودرو، چرخه تولید معکوس است؛ به این معنا که ابتدا محصول پیش فروش می‌شود و سپس تولید می‌شود. رمز موفقیت عاملین در صنعت خودرو، داشتن درک درست از زنجیره شرکت ها در این صنعت و یافتن حلقه‌هایی از این زنجیره است که عملکرد مجموعه آن‌ها با کسب مزیت رقابتی همراه است. بدین سبب داشتن استراتژی‌های مناسب، پیش شرطی برای کسب مزیت رقابتی بین شرکت‌های تولید خودرو است. امروزه استراتژی‌هایی که برای سازمان‌ها انتخاب می‌شود مبتنی بر رقابت است و موفقیت سازمان ها در گرو انتخاب به موقع استراتژی رقابتی است. در تدوین چنین استراتژی‌هایی، هدف، ایجاد ارزش بیشتر برای مشتری است. ایجاد چنین ارزشی منوط به مولفه هایی چون کیفیت، نوآوری، پاسخ گویی و کارا بودن شرکت هاست [۱]. مشتریان راضی از خدمت رسانی، منبع بقاء و سودآوری شرکت ها هستند. شرکت‌هایی که نمی‌توانند مشتریان خود را راضی نگه دارند، با گذشت زمان در بازار رقابتی نخواهند ماند. عرضه محصول با کیفیت برتر و ارایه خدمات به مشتریان در سطح عالی باعث ایجاد مزیت‌های رقابتی از جمله ایجاد موانع رقابتی، وفاداری مشتریان، تولید و عرضه محصولات متمایز، کاهش هزینه های بازاریابی و تعیین قیمت‌های بالاتر می‌شود[۲]. در صنعت خودرو، مشتری مداری به معنی پاسخگویی به نیاز مشتری در عملیات ارایه خدمات و همچنین در موارد بروز نقص فنی در کم‌ترین زمان ممکن است که با به حداقل رساندن مشکلات پیش آمده رضایت مشتری جلب شود[۳]. پیشگیری از ریزش مشتری، بخش مهمی از مدیریت ارتباط با مشتری، تلقی می‌گردد. از مهم‌ترین عوامل موثر بر ماندگاری مشتری، می‌توان به کیفیت خدمات ادراک شده، ارزش ادراک شده و رضایت مشتری اشاره کرد.

مشتریان عامل کلیدی و محوری در تقویت و ارتقاء سازمان ها قلمداد می‌شوند و جهت‌گیری کلیه اهداف، استراتژی‌ها و منابع حول محور جذب و نگهداری مشتری می‌باشد. اکثر صنایع به ویژه صنایع غیرقراردادی، که مراجعه مشتری به آنان دارای پراکندگی زمانی بسیاری می‌باشد، برای بررسی دقیق رفتار مشتریان خود از تکنیک‌ها و ابزار داده‌کاوی استفاده می‌نمایند. این امر موجب به‌دست‌آوردن اطلاعات دقیق‌تر شده که در نهایت منجر به ارایه سرویس و محصولات مطلوب به مشتری می‌گردد؛ علاوه بر این بررسی رفتار مشتری، می‌تواند به شناسایی جذب مشتریان جدید با صفات مشترک نیز کمک کرده و به جلوگیری از هزینه های تبلیغاتی گسترده بپردازد. در مطالعات گذشته عوامل متعدد موثر بر ماندگاری مشتری شناسایی شده‌اند که از مهم‌ترین آنها می‌توان به کیفیت خدمات ادراک شده، ارزش ادراک شده و رضایت مشتری اشاره کرد. در عصر جدید، توجه به رشد و توسعه ماندگاری مشتریان به عنوان عاملی مهم در استراتژی بازاریابی شرکت‌ها که منجر به حفظ مشتریان فعلی می‌شود، افزایش یافته است. امروزه مفهوم ماندگاری مشتری به عنوان دستورالعملی برای افزایش درآمد مطرح است. به دلیل وجود فشارهای رقابتی شدید، یکی از مهم‌ترین استراتژی‌هایی که سازمان‌های خدماتی از طریق آن می‌توانند به مزیت رقابتی پایدار دست یابند، بهبود کیفیت خدماتشان می باشد. به دلیل افزایش رقابت، سازمان ها می‌بایست به دنبال افزایش کیفیت خدمات خود بوده و در پروسه های زمانی مختلف اقدام به ارزیابی کیفیت

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

خدمات خود از دیدگاه مشتریان نمایند تا از این طریق از گردونه رقابت با دیگر سازمان ها عقب نمانند. به همین دلیل در این تحقیق سعی شده است تا با استفاده از روش‌شناسی استاندارد CRISP-DM و همچنین با تحلیل صحیح نیازهای مشتریان در دسته‌بندی‌های مختلف، مدلی برای شناسایی مشتریان رویگردان برای صنعت خودرو پیش‌بینی شود. لذا هدف از این تحقیق پیش بینی الگوی ماندگاری مشتریان در صنعت خودروساز با استفاده ازسیستم خبره می‌باشد. به طور کلی اهداف اصلی در انجام این تحقیق عبارت است از:

- شناسایی مشتریان با ارزش گروه با استفاده از معیار های مدل ارزش طول عمر مشتری در صنعت خودرو
- شناسایی مشتریان رویگردان و طبقه بندی مشتریان به دو گروه مانا و رویگردان در صنعت خودرو.

در ادامه این پژوهش، در بخش دوم ابتدا به بررسی مبانی نظری پرداخته و سپس به جهت مشخص کردن اهمیت موضوع و لزوم انجام پژوهش‌های بیشتر، به پژوهش‌های مرتبط پیشین اشاره خواهد شد. در بخش سوم روش پیشنهادی تشریح شده و در بخش چهارم نتایج شبیه سازی ارایه شده‌اند. همچنین نتیجه‌گیری در بخش پنجم بیان شده‌است.

۲ مبانی نظری و پیشینه پژوهش

کیفیت خدمات به عنوان یک اهرم استراتژیک، نقش ویژه‌ای در موفقیت سازمان‌های خدماتی دارد. بدین منظور بسیاری از صاحب نظران و محققان، مباحثی چون کیفیت خدمات، مدیریت کیفیت خدمات، ارزیابی کیفیت خدمات سازمان‌ها و دیگر مباحث را مورد توجه قرار داده‌اند. در دنیای رقابتی امروز، صنایع خودروساز برای حفظ مشتریان خود می‌بایست بیشترین توجه را به نیازها، نظرات و اعتراضات آنان در ارتباط با خدمات مربوط به محصول تولیدی خود داشته باشند. در این بین مواردی مانند تکریم مشتریان، تعامل مناسب با مشتریان، نظرسنجی از مشتریان، داشتن رابطه مناسب دوطرفه با مشتریان، بررسی مشکلات مشتریان و رسیدگی به موقع به شکایات آنان، آگاهی از رضایت‌مندی مشتریان، بالابردن کیفیت و سرعت ارایه خدمات، بررسی صحت خدمات ارایه‌شده به مشتریان و غیره می‌تواند کمک بسیار زیادی به شرکت‌های خودروساز در امر حفظ مشتریان نماید. امروزه نقش مشتریان از حالت پیروی از تولیدکننده، به هدایت تولیدکنندگان مبدل شده است، به همین دلیل دسته‌بندی مشتریان، در هدفمند ساختن سازمان‌ها در سفارشی‌سازی خدماتشان و نیز اولویت بندی محصولات بر اساس میزان سودآوری آن محصول کمک شایانی می‌کند[۴].

حفظ مشتری یکی از مهم‌ترین ابعاد مدیریت ارتباط با مشتری و به عبارت دیگر هسته‌ی اصلی آن به شمار می‌رود. ازجمله مفاهیم مرتبط با حفظ مشتری مفهوم رویگردانی مشتری است که انصراف مشتریان را از ادامه همکاری با شرکت موردتوجه قرار می‌دهد. رویگردانی مشتری موضوع مهمی برای همه‌ی سازمان‌هاست؛ زیرا امروزه در سازمان‌ها هزینه‌ای که صرف جذب مشتری جدید می‌شود، بسیار بیشتر از هزینه‌ای است که برای نگهداری مشتریان موجود انجام می‌شود. در این پژوهش، منظور از رویگردانی مشتری، نداشتن قرارداد در هیچ یک از ر شته خودروهای مختلف در سال جاری است.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

۲-۱ ارزش طول عمر مشتری

ارزش دوره عمر مشتری (CLV) یک سنجه‌ی مهم در بازاریابی و به‌طور خاص در مدیریت ارتباط با مشتری است[۵]. ارزش دوره عمر مشتری، میزان خالص ارزش فعلی به‌دست آمده از جریان پولی آینده، حاصل از یک مشتری در کل دوره‌ای که در سازمان فعال بوده و با آن تراکنش مالی دارد، می‌باشد. هنگامی که سازمان CLV را برای مشتریانش محاسبه می‌کند، می‌تواند به صورت بهینه منابع محدود خود را برای به‌دست آوردن بیشترین برگشت تخصیص دهد. CLV می‌تواند به عنوان معیاری برای راهنمایی تخصیص منابع در یک سازمان مشتری محور، مورد توجه قرار گیرد.

۲-۱-۱ انتخاب مشتری

تحقیقات نشان داده است که مشتریان وفادار لزوماً سودآور نیستند. سود ناشی از بسیاری از مشتریان، بسیار کمتر از هزینه‌ای است که سازمان باید برای نگهداری آنها داشته باشد. انتخاب مشتریان مناسب برای نگهداری-کسانی که بیشترین سود را برای سازمان داشته باشند-گام مهمی در بهبود سودآوری است. تحقیقات نشان داده است که سازمانها نیاز به معیاری برای اندازه‌گیری سودآوری هر مشتری دارند تا در خصوص انتخاب مشتریان تصمیم بگیرند. محاسبه ی CLV که بیانگر سودآوری مشتری در آینده خواهد بود، برای این امر بسیار مناسب است. رفتارهای متفاوت مشتریان، کلید مدیریت ارتباط با مشتری سودآور است. با بخش‌بندی مشتریان بر اساس CLV آنها، می‌توان برای هر بخش استراتژی‌های به‌خصوصی را ایجاد کرد. در بسیاری از موارد بودجه سازمان‌ها محدود است و نمی‌توانند منابع خود را به همه ی مشتریانشان تخصیص دهند. در این صورت شناسایی مشتریان با ارزش و تخصیص بودجه برای آنان، به جهت سودآوری سازمان از اهمیت بالایی برخوردار است.

۲-۲ مدل پایه CLV

در این مدل، درآمد در هر دوره از میزان هزینه آن دوره کسرشده و بر ضریب نرخ بهره تقسیم شده است ، تا سود خالص مراودات و تراکنش‌های مالی مشتری با سازمان در آن دوره به‌دست آید. به منظور محاسبه جریان درآمدی در بین هر دوره مقدار ۰/۵ از هر دوره کاسته شده است. در این مدل فرض شده است که همه‌ی جریان نقدی در انتهای دوره رخ می‌دهد.

$$CLV=\sum_{i=0}^{n}\frac{(R_i-C_i)}{(1+d)^{i-0/5}} \qquad (1)$$

R_i : میزان درآمد سازمان از مشتری ، C_i: میزان هزینه‌ای که سازمان صرف محصول یا خدمت می‌کند،

d : ضریب نرخ بهره، i : دوره جریان نقدینگی حاصل از تراکنشات مشتری، n : تعداد دوره‌ها

این مدل ساده‌ترین مدل ارزش طول عمر مشتری است. از ویژگی‌های آن می‌توان به در نظرگرفتن دوره ثابت گردش مالی در هر دوره زمانی، لحاظ نمودن مشتریان حاضر در سازمان، نادیده‌گرفتن هزینه‌های کسب مشتری و همچنین در نظر نگرفتن خرید احتمالی مشتریان و زمان جریان نقدینگی اشاره کرد[۶].

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

گلیدی و همکاران در[۷] تعریف جدیدی از رویگردانی را معرفی می‌کنند. آنان رویگردان را کسی می‌دانند که CLV وی در طول زمان در حال کاهش باشد. آن‌ها معتقدند که CLV مشتریان قبل و بعد از کمپین نگهداری تغییر می‌کند. همچنین آنان از CLV برای تعریف یک معیار عملکردی جدید استفاده کرده‌اند.

همان‌طور که از محاسبه‌ی CLV مشاهده می‌شود، ارزش دوره‌ی عمر مشتری تا حد زیادی به این موضوع وابسته است که مشتری در دوره‌ی بعدی زمانی در سازمان فعال است، حضور دارد یا خیر. این موضوع به طور خاص در صنایع غیرقراردادی مهم است؛ زیرا مشتری آزادی دارد که هر زمان که بخواهد رابطه‌ی خود را با سازمان قطع کند. در نتیجه احتمال ماندگاری مشتری تأثیر زیادی بر CLV دارد.

۲-۳ پیشینه تحقیق

با توجه به اهمیت بالای وفاداری مشتریان در میزان موفقیت سازمان‌ها، در برخی از پژوهش‌ها به کاربری داده کاوی در تشخیص رویگردانی مشتریان پرداخته شده است.

در[۸] از الگوریتم‌های داده‌کاوی جهت ساخت مدلی به منظور پیش‌بینی رویگردانی مشتریانی که از کارت اعتباری استفاده می‌کنند، استفاده شده است. این تحقیق بر روی پایگاه داده مشتریان یک بانک چینی صورت گرفته و از چهار دسته متغیر اطلاعات مشتری، اطلاعات کارت اعتباری، داده‌های مربوط به ریسک مشتری و اطلاعات مربوط به تراکنش‌ها استفاده شده که در مجموع شامل ۱۳۵ متغیر می‌باشد. در این تحقیق مشتری رویگردان فردی تعریف شده که در طول دوره مشاهده دوازده ماهه هیچ تعاملی را با بانک نداشته است. در این تحقیق الگوریتم‌های رگرسیون و درخت تصمیم جهت دسته‌بندی مشتریان انتخاب شده و نتایج تحقیق نشان داد که عملکرد مدل‌های رگرسیونی کمی بهتر از عملکرد درخت‌های تصمیم بوده است. زی در [۹] با تاکید توجه بانک‌های چینی به رویگردانی مشتری و تلاش برای به حداکثر رساندن رضایت مشتریان، از یک روش یادگیری جدید به نام جنگل‌های تصادفی متوازن بهبودیافته (IBRF) جهت تعدیل مشکل بحث برانگیز عدم توازن در توزیع داده‌ها در تحلیل رویگردانی مشتری استفاده کرده‌است. این روش با اعمال بر روی داده‌های واقعی مربوط به رویگردانی مشتری یک بانک چینی، به طور قابل توجهی دقت پیش‌بینی را در مقایسه با سایر الگوریتم‌ها مانند شبکه‌های عصبی مصنوعی، درخت‌های تصمیم و ماشین بردار پشتیبان بالا برده‌است. این روش در مقایسه با سایر الگوریتم‌های جنگل تصادفی مانند الگوریتم جنگل تصادفی متوازن و الگوریتم جنگل تصادفی وزن‌دار بهتر عمل کرده است. لین بیان می‌کند که از منظر اقتصاد و مدیریت ریسک، درک ویژگی‌های مشتری جهت حفظ مشتری ضروری به نظر می‌رسد و لازم است سازمان‌ها بین مشتریان معتبر و مشتریان بد، تمایز قایل شوند. این مقاله تئوری مجموعه‌های ناهموار را که یک روش تصمیم‌گیری بر اساس قوانین است، برای استخراج قوانین مرتبط با رویگردانی مشتری بکار می‌گیرد. سپس از گراف شبکه جریان که یک رویکرد وابسته به مسیر است برای کشف قوانین و متغیرهای تصمیم‌گیری استفاده می‌نماید و در نهایت ارتباط بین قوانین و انواع مختلف روگردانی را نشان می‌دهد. این تحقیق بر روی مجموعه‌ای ۲۱۰۰۰ نفره از مشتریان صورت گرفته و آن‌ها را به سه دسته مشتریان پایدار، رویگردانی داوطلبانه و رویگردان غیرداوطلبانه تقسیم می‌کند. نتایج نشان داد که این مدل ترکیبی

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

پیش‌گویی خوبی از روی‌گردانی مشتری داشته و اطلاعات مفیدی برای تصمیم گیران در تدوین استراتژی‌های بازاریابی فراهم می‌کند[۱۰].

دیجیتالی شدن، سازندگان خودرو را به سمت تغییر ارزش پیشنهادی خود سوق می‌دهد و به سمت همکاری و ارزش‌گذاری بیشتر برای مشتری هدایت می‌کند. تغییر به سمت ارایه خدمات، مستلزم تغییر تحولی از محصول محوری به مشتری محوری است. در [۱۱] یک چارچوب مرجع مفهومی (CRF) از دیدگاه مدل کسب‌وکار برای سیستم‌بندی سیستم‌های خدمات خودرو پیشنهاد شده است. CRF ابعاد و وابستگی‌های مربوطه را بین ذینفعان درگیر و زیرساخت های لازم به منظور تسهیل مفهوم سازی خدمات دیجیتال را درمراحل اولیه طراحی خدمات ارائه می دهد. نتایج نشان می‌دهد که ایجاد ارزش برای خدمات خودرو در شبکه‌های مشترک بین سهامداران وابسته به یکدیگر رخ می‌دهد که در آن مشتریان نقشی جدایی‌ناپذیر در طول چرخه عمر خدمات بازی می‌کنند. ارزیابی کیفیت خدمات به طور گسترده ای در خدمات پس از فروش به ویژه در صنعت خودرو استفاده شده است. هدف مطالعه [۱۲] تعیین عوامل موثر بر رضایت مشتری درخدمات پس از فروش خودرو در تویوتا Dasmarinas-Cavite فیلیپین با استفاده از رویکرد SERVQUAL بود. چندین بعد SERVQUAL مانند موارد قابلیت اطمینان، پاسخگویی، اطمینان و همدلی به طور هم‌زمان برای رضایت مشتری تجزیه و تحلیل شد. مدل‌سازی معادلات ساختاری (SEM) نشان داد که از بین پنج بعد SERVQUAL، قابلیت اطمینان و همدلی با رضایت مشتریان در شرکت ارتباط معنی‌داری دارند.

در [۱۳] رضایت مشتری به صورت تابعی خطی فرموله‌شده تا رضایت مشتری از شرکت های خودروسازی ایران بهینه‌سازی شود. با استفاده از روش AHP پرسشنامه خبره، رضایت مشتری فرموله شد. همچنین با در نظر گرفتن شاخص‌هایی به عنوان متغیرهای کنترلی، رضایت مشتری به عنوان متغیر وضعیت، مدل‌سازی‌شده و مدل کنترلی بهینه ارایه شده است. این مدل رضایت مشتری بهینه را برای هر سال با توجه به نوع خودرو و یک دوره خاص محاسبه می کند. شرکت‌های ایران خودرو و سایپا از مهم‌ترین تولیدکنندگان کشور می باشند و درصد بالایی از مشتریان داخل کشور را به خود اختصاص می‌دهند. هدف از انجام تحقیق [۱۴]، سنجش و ارزیابی رابطه بین خدمات پس از فروش این دو شرکت و میزان رضایت مشتریان می‌باشد. سه شاخص اساسی برای ارزیابی خدمات پس از فروش انتخاب شد که عبارتند از: خدمات نمایندگی‌های مجاز، تامین قطعات و تضمین. با تدوین فرضیاتی بر اساس این شاخص‌ها، داده‌های جمع‌آوری شده به وسیله پرسشنامه با استفاده از نرم افزار SPSS برای هر نوع خودرو این شرکتها به طور جداگانه مورد آزمون قرار گرفت. مقاله [۱۵] با تحقیق بر روی استراتژی های پایدار شرکت‌های پیشرو در صنعت خودرو، میزان برآورده شدن برخی از اهداف SDG سازمان ملل توسط این شرکت‌ها در جهان را بررسی کرده است. روش تحلیل تطبیقی کیفی با استفاده از روش جایگزین میز اطلاعاتی را در مورد استراتژی‌های پایداری و تحقق آنها در بین پنج شرکت برتر صنعت خودرو بر اساس مقیاس اینتربرند از سال ۲۰۲۱ با در نظر گرفتن اهداف پایداری دستور کار ۲۰۳۰ سازمان ملل ارایه کرد. نتایج تجزیه و تحلیل نشان داد که شرکت‌های مشاهده‌شده در بخش خودرو اکثر اهداف توسعه پایدار سازمان ملل را با استراتژی‌های پایداری خود برآورده می کنند که نشان می دهد آنها با موفقیت عملیات تجاری خود را با الزامات تجاری مدرن

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

تطبیق می‌دهند. این همچنین باعث سبز شدن برندهای آنها می‌شود، از این رو استفاده از اصل پایداری نیز بر قدرت برندهای آنها تثیر می گذارد و "ارزش مشتری گسترده" را اضافه می‌کند. در مطالعه [۱۶] برای کمک به بازار پس از فروش خودروهای اسپانیایی، با استفاده از مدل‌های رگرسیون، مهم‌ترین متغیرها برای بخش خودرو برآورد شدند که بر پیش‌بینی‌های بلندمدت درآمدهای پس‌بازار خودرو تاثیر می‌گذارند. سپس، روش‌های مشارکتی برای تعیین کمیت تاثیر عوامل شرطی‌سازی جدید اعمال شدند. نتایج نشان داد که ادراک ذینفعان، پیش‌بینی‌ها را برای بخش‌های اقتصادی درگیر در یک مدل کسب‌وکار در حال تغییر تعدیل می‌کند. پژوهش [۱۷] رویکردی را برای ارزیابی عوامل مؤثر بر رابطه فروشنده-تولیدکننده و ایجاد راه حلی برای پیش‌بینی چگونگی ادامه یا پایان همکاری تولیدکنندگان با نمایندگی‌های خود در شبکه خدمات پس از فروش ارایه کرده است. برای انجام این‌کار، از ابزارهای تحلیل و روش‌های طبقه‌بندی برای پیش‌بینی همکاری فروشندگان به‌عنوان یک متغیر وابسته در دو کلاس بر اساس داده‌های جمع‌آوری‌شده دریک دوره ۲ ساله استفاده شد. نتایج نشان داد که نقض عملکرد نمایندگی، تغییر اصول نمایندگی و درآمد خدمات بیشترین تاثیر را بر همکاری نمایندگی‌ها در شبکه خدمات پس از فروش خودرو دارد. همچنین، در این تحقیق یک راه حل مدیریت نمایندگی توسط تولیدکنندگان پیشنهاد شد که می تواند در موارد واقعی در شبکه های خدمات پس از فروش برای جلوگیری از استفاده از منابع اضافی، بهبود عملکرد ضعیف نمایندگی‌ها، جلوگیری از قطع بی‌مورد همکاری، جلوگیری از هزینه‌های جذب نمایندگی‌های جدید و کاهش نارضایتی مشتریان استفاده شود.

۳ روش پیشنهادی

در این بخش اجرای روش پیشنهادی گام به گام تشریح می‌شود. این تحقیق در سه مرحله‌ی اصلی به‌دست آوردن احتمال رویگردانی مشتریان ، محاسبه‌ی ارزش دوره عمر مشتریان و انتخاب مشتریان برای تخصیص بهینه‌ی منابع بازاریابی انجام شده است. در این تحقیق مشتری رویگردان مشتری است که در سال ۱۳۹۶ به عنوان مشتری در شرکت خودرو حضور داشته باشد و در سال ۱۳۹۷ در هیچ رشته‌ی خودروای، قراردادی با شرکت نداشته باشد. براین اساس وضعیت رویگردانی مشتریان بررسی شد و برای هر کدام ثبت گردید.

۳-۱ جمع‌آوری داده

داده‌های مورد بررسی در این تحقیق، داده‌های خودرونامه‌های صادرشده در شرکت ایران خودرو است. این داده‌ها به صورت فایل‌های متنی با پسوند CSV، در چند مرحله و به تفکیک سال صدور و دسته‌بندی‌های خودرو در دسترس قرار گرفته است. داده‌های خام موجود در پایگاه داده اغلب به صورت پردازش نشده و نویزی هستند. در داده‌های دریافتی، داده‌های گم شده و زاید وجود داشت که در مرحله پاکسازی داده باید حذف شوند. همچنین با توجه به آن‌که موضوع تحقیق بر روی مشتریان چندقراردادی است، مشتریان تک قراردادی باید شناسایی و حذف شوند. با توجه به اینکه داده‌ها برای رشته خودروها و سال‌های مختلف، به صورت فایل‌های متنی مجزا دریافت شده بود. در مرحله یکپارچه سازی باید همه‌ی داده‌ها در یک پایگاه داده قرار بگیرد. همچنین

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

جدول جدید برای هر کدام از مشتریان تشکیل شود. برای آنکه ایجاد مشخصه‌های مورد نیاز داده‌کاوی ممکن شود، باید از میان جداول مختلف، جدولی ایجاد می‌شد که در آن مشخصات و داده‌های مربوط به هر مشتری وجود داشت و هر رکورد آن مربوط به یک مشتری می‌بود. پس از آن با توجه به ادبیات رویگردانی مشتری و نیز داده‌های موجود، مشخصه‌های لازم ساخته شدند. در این مرحله علاوه بر تعیین مشخصه‌های تصمیم‌گیری، مشخصه‌ی هدف که رویگردانی مشتری است نیز برای مورد مطالعه تعریف گردید. همچنین، به بررسی مدل‌های مورد استفاده پیشین پرداخته شد که بر این اساس سه مدل آموزش یافته به روش‌های جنگل‌های تصادفی، درخت‌های فزاینده و مدل ترکیبی (Hybrid) با انتخاب منطقی مورد بررسی قرار گرفته و میزان دقت هر مدل بر اساس مجموعه داده‌های مورد استفاده تعریف شد [۲۱-۱۸]. داده‌های مورد استفاده در این مطالعه بر اساس روش Hold Outبه صورت تصادفی با نرخ ۷۰ به ۳۰ به مجموعه داده‌های آموزش و آزمایش تقسیم شده و سپس مدل بر اساس آن‌ها پیاده‌سازی شد.

۳-۲ ارزیابی الگو

پس از آن‌که مدل پیش‌بینی‌کننده ایجاد شد، از آن برای پیش‌بینی رفتار آینده‌ی مشتریان استفاده می‌شود. یکی از گام‌های مهم، اطمینان از صحت عملکرد مدل پیش‌بینی است.

۳-۲-۱ نمودار lift و Gain

معروف‌ترین راه برای مقایسه‌ی عملکرد مدل‌های دسته‌بندی، استفاده از نسبتی به نام lift است. آن‌چه lift در واقع سنجش می‌کند، تغییر در چگالی یک کلاس خاص هنگامی که مدل برای انتخاب گروهی از جامعه‌ی کل استفاده می‌شود، می‌باشد [۲۲].

$$Lift = P(class\ t \mid sample) \ / P(class\ t \mid population) \qquad (۲)$$

lift تابعی از سایز نمونه است. زمانی که lift محاسبه می‌شود، نرم‌افزار ابتدا رکوردها را برحسب احتمال این‌که آن‌ها به عنوان مثبت دسته‌بندی شوند، مرتب می‌کند. پس از آن Lift به ازای درصد مختلف اندازه نمونه با ترتیبی که در ادامه ذکر می‌گردد، رسم می‌شود: مقادیر این نرخ از چپ به راست کاهش یافته و در ۱۰۰ ٪ اندازه نمونه به مقدار یک می‌رسد. هر چه این نرخ بالاتر باشد، بیانگر دقت بهتر مدل کلاس‌بندی است. به طور شهودی نرخ چگالی در ۱۰ ٪ بالای جامعه بایستی بیشتر از دو برابر این نرخ در کل جامعه باشد.

نمودارهای lift و gain، روشهای اعتبارسنجی گرافیکی از تخمین و مقایسه عملکرد وکارایی مدل‌های دسته‌بندی هستند. نمودارهای lift معمولاً در شکل تجمعی نشان داده می‌شوند که اصطلاحاً به این فرم نمودار gainگفته می‌شود. خط قطری کشیده شده در نمودار gain برابر با محور افقی ۱ = lift در نمودار lift است. معمولاً نمودارهای gain به آرامی از این خط قطری فاصله گرفته و نموداری منحنی شکل را در بالای قطر اصلی تا رسیدن دوباره به آن طی می‌کنند.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

۳-۳ محاسبه‌ی CLV مشتریان

پس از به‌دست آوردن احتمال رویگردانی مشتریان در مرحله‌ی قبل، در این مرحله ارزش طول عمر مشتریان محاسبه خواهد شد. در این تحقیق برای CLV مانند [۷] از تعریف ارایه‌شده توسط [۲۳] استفاده می‌شود. ارزش طول عمر مشتری به صورت مجموع دریافت‌های آینده مورد انتظار، هنگامی که مشتری سودی برای هر دوره تولید کند، تعریف می‌شود. همچنین با توجه به توضیح [۷] ، چون بر روی نگهداری تمرکز شده است نه جذب مشتری، همه‌ی مشتریان در گذشته جذب شده‌اند و تنها سود دریافتی از آن‌ها محاسبه می‌شود و هزینه‌های جذب مشتری و هزینه‌های ثابت دیگر در نظر گرفته نمی‌شود. در نتیجه فرمول CLV مورد استفاده به صورت زیر خواهد بود:

$$CLV = \sum_{t=0}^{T} \frac{(m_t r_t)}{(1+i)^t} \quad (۳)$$

r_t : احتمال خرید مجدد مشتری یا فعال بودن وی در زمان t ، i : نرخ تنزیل ، T: افق زمانی برای تعیین CLV برای محاسبه‌ی سود مورد انتظار از مشتری، از تاریخچه‌ی مالی در سال‌های گذشته استفاده می‌شود. سپس با استفاده از رگرسیون خطی درآمد دوره‌ی بعدی مشتری محاسبه می‌شود. همچنین نرخ تنزیل، نرخ تورم اعلام شده از سوی بانک مرکزی خواهد بود.

۳-٤ مدل RFM

یکی از مدل‌های مطرح در تحلیل ارزش مشتری، مدل RFM است که توسط Hughes در سال ۱۹۹۴ ارایه‌شده است و به بیان تفاوت مشتریان با استفاده از سه متغیر تازگی، تکرار و ارزش پولی می‌پردازد[۲۴]. پارامترهای مدل RFM به شرح زیر است:

- تازگی آخرین خرید (R): نشان‌دهنده مدت زمان بین آخرین تعامل تجاری با زمان حال می‌باشد، هر چه این مدت کمتر باشد، R بیشتر است.
- تکرار خرید (F): تعداد تکرار تراکنش‌ها را در یک بازه مشخص نشان می‌دهد، برای مثال دو بار در سال، دو بار در یک فصل یا دو بار در یک ماه. هر چه تکرار بیشترباشد، F بزرگ‌تراست.
- ارزش پولی خرید (M): ارزش پولی تعاملات را در بازه‌های خاص نشان می‌دهد. هرچه ارزش پولی بیشتر باشد، M بزرگ‌تراست.

تحقیقات نشان می‌دهد که هر چه R و F بیشتر باشند، احتمال آنکه تراکنش جدیدی با مشتری صورت بگیرد، بیشتر است[۲۴]. همچنین اگر M بزرگ‌تر باشد، بازگشت مشتری برای خرید بیشتر است. متغیرهای مدل RFMبرای بخش‌بندی مشتریان بسیار کارا هستند. از دیگر کاربردهای مدل RFM در بخش‌بندی مشتریان بر اساس متغیرهای تازگی، تکرار و ارزش پولی، می‌توان به بخش‌بندی مشتریان به منظور تعیین سیاست‌های بهینه بازاریابی برای هر بخش اشاره کرد.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

٤ پیاده‌سازی

در این بخش مراحل اجرای روش پیشنهادی گام به گام توضیح داده می‌شود. مرحله‌ی اول در مدل پیشنهادی یافتن احتمال رویگردانی مشتریان است. برای این منظور همانطور که قبلاً اشاره شد، از ابزارهای داده‌کاوی استفاده می‌شود. در این پژوهش، شناسایی مشتریان با استفاده از پارامترهای مورد نظر برای نیل به اهداف آن، به عنوان یکی از نیازهای اساسی صنعت خودرو مطرح گردیده است.

٤-١ مدل ترکیبی پیش‌بینی ماندگاری مشتریان

در شکل ١ مدل‌های استاندارد نشان داده شده است. مدیریت رویگردانی در سه گام اصلی به صورت زیر انجام می‌شود: ١ - مدل تعیین مشتریان با ارزش، ٢ - مدل پیش بینی رویگردانی، ٣ - مدل نگهداری مشتریان.

٤-٢ جمع‌آوری و آماده‌سازی و انتخاب داده‌ها

در این مرحله ابتدا به بررسی کامل پایگاه داده‌های سازمان به منظور شناخت دقیق جداول، فیلدها و روابط بین جداول پرداخته شده است. سازمان مورد مطالعه یک شرکت خودرو در ایران است که دارای شعب متعددی در سراسر ایران می‌باشد. در مرحله بعد، فیلدهای مورد نیاز از پایگاه داده سازمان استخراج گردید. فایل‌های متنی پس از دریافت به جداول پایگاه‌داده در نرم‌افزار SQL SERVER وارد شدند. تنها داده‌ی با ارزشی که می‌توان مشتری را با آن شناخت و برای هر مشتری یکتاست، کد ملی وی است. در این تحقیق مشتری رویگردان به مشتری‌ای اطلاق می‌شود که در سال ١٣٩٦ به عنوان مشتری در شرکت خودرو حضور داشته باشد و در سال ١٣٩٧ در هیچ رشته‌ی خودرویی ، قراردادی با شرکت نداشته باشد. بر این اساس وضعیت رویگردانی مشتریان بررسی و ثبت گردید.

انتخاب مشخصه، فرایند شناسایی فیلدهای داده‌ای است که بهترین نتیجه را برای پیش‌بینی حاصل می‌کنند [٢٥]. به‌طور معمول برای انتخاب مشخصه‌های مناسب، از نظر کارشناسان مربوطه و همچنین مشخصه‌های بکار رفته در ادبیات استفاده می‌شود. در این تحقیق با توجه به داده‌های موجود، از اطلاعات تراکنشی و اطلاعات دموگرافیک مشتریان استفاده شده است. همچنین در انتخاب مشخصه‌ها به چند قراردادی بودن مشتریان نیز توجه شده است.

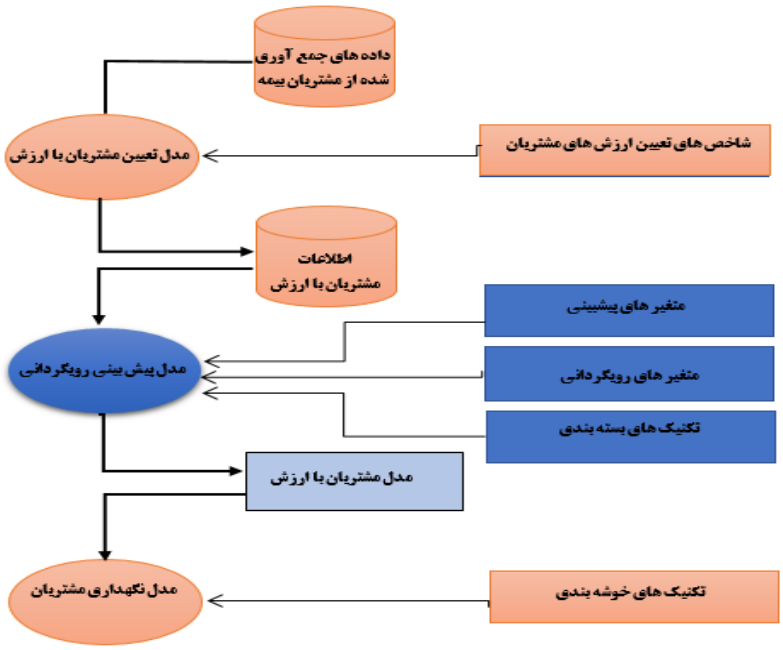
٤-٣ ساخت مدل‌های داده‌کاوی

از آنجایی که داده‌های مورد بررسی نامتوازن هستند و نسبت مشتریان رویگردان به مشتریان غیررویگردان بسیار کم است، باید از یکی از روش‌های متعادل‌ساز استفاده شود. برای این امر از روش متعادل‌سازی با جایگذاری برای نمونه‌گیری از کلاس کمیاب (رویگردان) استفاده شد و نسبت کلاس رویگردان به غیر رویگردان ١ به ٢ تنظیم شد. در نتیجه تعداد کل نمونه‌ها به ٥٧١١ رسید که از این تعداد، ١٩٢٤ نمونه مربوط به مشتریان رویگردان است.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

در این تحقیق از روش‌های درخت تصمیم CHAID، شبکه عصبی پرسپترون و K نزدیکترین همسایه، استفاده شده است. در استفاده از روش k نزدیکترین همسایه، این روش با مقادیر k از ۱ تا ۱۰ اجرا شد تا بهترین جواب پیدا شود.



شکل ۱. چارچوب استاندارد برای پیش‌بینی رویگردانی مشتری

٤-٤ اعتبارسنجی روش‌های داده‌کاوی مورد استفاده

در این تحقیق از دو معیار دقت کلی و Lift به عنوان معیارهای اصلی ارزیابی ومقایسه‌ی روش‌های دسته‌بندی برای ارزیابی روش‌ها استفاده شده است.

٤-٤-١ ماتریس ارزیابی

برای ارزیابی عملکرد مدل‌های دسته‌بندی با متغیر هدف دودویی، ماتریس انطباق مورد استفاده قرار می‌گیرد[۲۶]. بر مبنای ماتریس انطباق، سه شاخص مهم شامل حساسیت، تشخیص و صحت به صورت زیر تعریف می‌شوند:

$$Sensitivity = TP / (TP + FN) \quad (۴)$$

$$Specificity = TP / (TP + FP) \quad (۵)$$

$$Accuracy = TP + TN / (TP + FP + TN + FN) \quad (۶)$$

در این تحقیق از نرخ خطای کلی و نمودار lift برای اعتبارسنجی نتایج استفاده شده است.

دقت کلی: نسبت نمونه‌هایی که به درستی تشخیص داده شده‌اند به کل نمونه‌ها

منحنی Lift : پس از مرتب‌سازی نمونه‌ها بر حسب نرخ اطمینان، نسبت پارامتر مورد پیش‌بینی در بخش‌های مختلف (مثال ۲۰ ٪ اول) به کل جامعه، مقدار Lift را نشان می‌دهد.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

جدول۱ مقادیر به‌دست آمده برای این معیارها را در الگوریتم‌های مورد استفاده نشان می‌دهد. همان‌طور که مشاهده می‌شود روش k نزدیک‌ترین فاصله بهترین جواب را هم از نظر دقت کلی و هم از نظر معیار lift دارد. این روش ابتدا برای مقادیر مختلف k انجام گرفت، پس از بررسی دقت کلی به ازای مقادیر مختلف، مشاهده شد که بهترین جواب در k=۱ به‌دست آمده است. از این‌رو نتایج برای k=۱ به عنوان نتیجه‌ی روش k نزدیک‌ترین همسایه در نظر گرفته شده است . همچنین جدول ۲ ماتریس انطباق را برای روش k نزدیکترین همسایه نشان می‌دهد.

جدول ۱. مقایسه نتایج روش‌های دسته‌بندی

روش	دقت کلی	Lift ۲۰٪
درخت تصمیم	٪۸۳/۷۵	۲/۱
شبکه عصبی	٪۷۰/۸۹	۱/۸
K نزدیکترین همسایه	٪۹۷/۴۳	۲/۹

جدول ۲. ماتریس انطباق روش k نزدیکترین همسایه

	F	T
F	۳۷۶۲	۲۴
T	۵۲	۱۸۷۲

نرخ پارامترهای حساسیت و تشخیص به صورت زیر است:

$$\text{حساسیت} = \frac{TP}{(TP+FN)} = \frac{۱۸۷۲}{(۱۸۷۲+۵۲)} = ۰/۹۷$$

حساسیت = TP / (TP+FN) = ۱۸۷۲ / (۱۸۷۲ + ۵۲) = ۰/۹۷

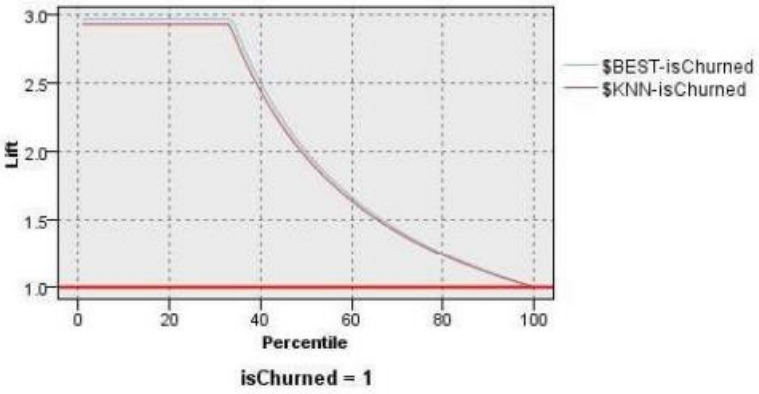
تشخیص = TP / (TP+FN) = ۳۷۶۲ / (۳۷۶۲ + ۲۴) = ۰/۹۹

پارامتر حساسیت نشان‌دهنده‌ی آن است که به چه نسبتی عناصر مثبت نسبت به مقدار واقعی خود پیش‌بینی شده‌اند که در پژوهش حاضر، نسبتی از رویگردانان واقعی را نشان می‌دهد که از کل جامعه‌ی رویگردانان توسط مدل پیش‌بینی‌کننده شناسایی شده‌اند. به عبارت دیگر، مدل ۹۷ درصد رویگردانان را شناسایی کرده است. پارامتر تشخیص بیانگر نسبت برای عناصر منفی جامعه است. این پارامتر بیانگر آن است که چه درصدی از افراد غیر رویگردان واقعی توسط مدل پیش‌بینی‌کننده درست شناسایی شده‌اند. مدل ۹۹ درصد افراد غیر رویگردان را شناسایی کرده است.

شکل ۲ نمودار lift روش k نزدیک‌ترین همسایه را نشان می‌دهد. به طور معمول اگر نرخ lift به ازای ۱۰% اول از ۲ بیشتر باشد، آن مدل دسته‌بندی در جامعه‌ی مورد بررسی قابل قبول است. با توجه به شکل مشاهده می‌شود این مقدار برای مدل مورد استفاده بالاتر از ۵.۲ است که این موضوع اعتبار مدل را در دسته‌بندی مشتریان نشان می‌دهد.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

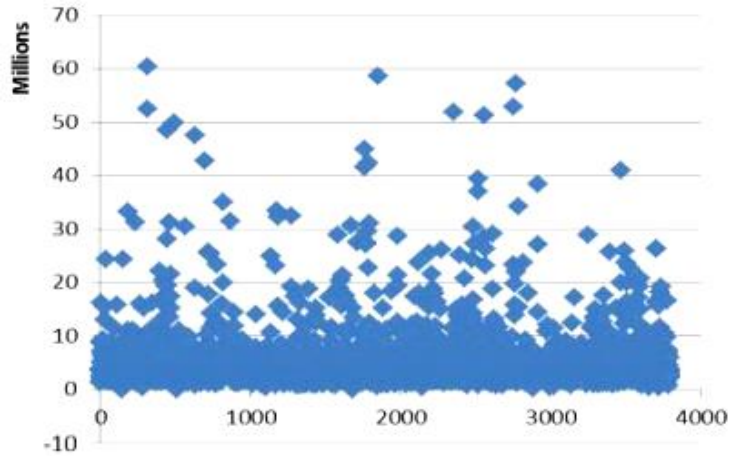
[DOI: 10.71773/jamlu-2025-1-2127]



شکل ۲. نمودار lift روش k نزدیکترین همسایه

٤-٥ محاسبه ارزش عمر مشتری

پس از به‌دست آوردن احتمال رویگردانی مشتریان می‌بایست به کمک آن برای مشتریان CLV آینده محاسبه شود. در این تحقیق ارزش طول عمر مشتری برای یک دوره‌ی بعد پیش‌بینی می‌شود. برای محاسبه‌ی CLV، سال ۱۳۹۶ را سال مبنا قرارداده و CLV برای سال ۱۳۹۷ پیش‌بینی می‌شود. برای محاسبه‌ی درآمد سال آینده‌ی مشتری (m) ، از مدل رگرسیون خطی استفاده می‌شود. در این مدل مبلغ درآمد حاصل از مشتری، متغیر پاسخ و زمان به عنوان متغیر پیش‌گو منظور می‌شود. پس از به‌دست آوردن مقدار برای سایر مشتریان، مقدار معیار r^2 برای هر کدام محاسبه شد. میانگین r^2 برابر ۰/۷۵ است. نرخ تنزیل در معادله‌ی ۳ برابر با نرخ تورم اعلام شده از بانک مرکزی در نظر گرفته می‌شود که این نرخ برای سال ۱۳۹۷ برابر ۳۰/۵ است. شکل۳ پراکندگی مقدار CLV مشتریان را نشان می‌دهد.



شکل ۳. پراکندگی ارزش دوره عمر مشتریان

٤-٦ تشخیص تعداد بهینه خوشه‌های مشتریان

یکی از مسایل مهم در خوشه‌بندی، تعیین تعداد بهینه خوشه‌ها است که در اکثر الگوریتم‌ها مانند k-means باید توسط خود کاربر معین شود. یک راه ممکن و رایج، جهت تعیین تعداد خوشه بهینه، امتحان kهای مختلف و

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

تعیین مقدار بهینه براساس یک سری شاخص‌های از پیش تعریف شده است. در این تحقیق جهت مشخص کردن تعداد خوشه‌های بهینه از روش شاخص دون Dunn (Index) استفاده شده است. این معیار توسط رابطه (۷) تعریف می‌شود:

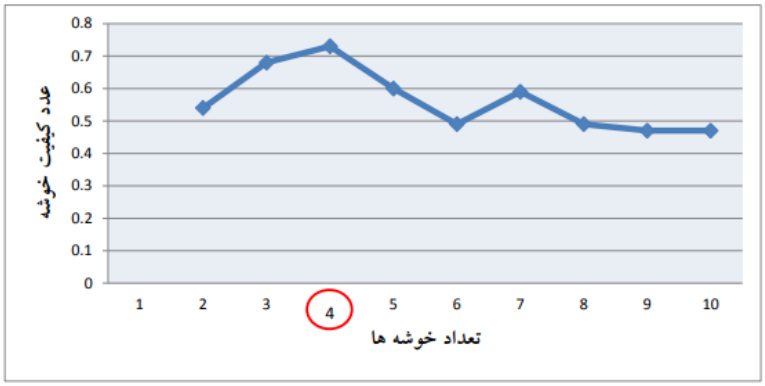
$$D=\min_{i=1\dots n_c}\left\{\min_{j=i+1\dots n_c}\left[\frac{d(c_i,c_j)}{\max_{k=1\dots n_c}(diam(c_k))}\right]\right\} \quad (۷)$$

که $d(c_i,c_j)$ و $diam(c_i)$ در آن به ترتیب با روابط (۸) و (۹) محاسبه می‌شوند.

$$d(c_i,c_j)=min_{x\in c_i\ ,y\in c_j}\{d(x,y)\} \quad (۸)$$

$$diam(c_i)=\ max_{x,y\in c_i}\{d(x,y)\} \quad (۹)$$

اگر مجموعه داده‌ای، دارای خوشه‌هایی جداپذیر باشد، انتظار می‌رود فاصله بین خوشه‌ها زیاد و قطر خوشه‌های (Diameter) آن کوچک باشد. در نتیجه مقداری بزرگ‌تر برای رابطه این معیار، مقداری مطلوب‌تر است.



شکل ٤. نمودار تعداد بهینه خوشه‌ها

ابتدا با الگوریتم K- means، خوشه بندی برای ۲ تا ۱۱ خوشه انجام گرفت و شکل ۴ نشان می‌دهد که تعداد بهینه خوشه ها عدد ۴ است. زیرا هر چه مقدار به دست آمده از شاخص دان بزرگ‌تر باشد، بهتر است و تعداد خوشه‌ای که مقدار این شاخص را زیادتر کند، مقدار بهینه تعداد خوشه‌ها است. جدول ۳ توضیحات کامل‌تری را درباره شاخص دان(Duun) نشان می دهد.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

جدول ۳. مقادیر شاخص دان(Duun) برای خوشه‌بندیK- means

	تعداد خوشه=۲	تعداد خوشه=۳	تعداد خوشه=۴	تعداد خوشه=۵	تعداد خوشه=۶	تعداد خوشه=۷	تعداد خوشه=۸	تعداد خوشه=۹	تعداد خوشه=۱۰
خوشه شماره ۱	۱/۰۲۲*	۳/۸۵۲	۳/۹۴۱	۴/۰۳۲	۴/۰۳۲	۳/۸۹۵	۳/۴۹۰	۴/۸۷۰	۴/۹۵۰
خوشه شماره ۲	۱/۰۸۰	۱/۱۷۷*	۳/۴۸۲*	۴/۹۱۰	۴/۹۱۰	۳/۸۹۰	۴/۵۶۰	۴/۴۴۰	۴/۴۴۵
خوشه شماره ۳		۳/۶۵۰	۳/۷۴۲	۴/۸۱۶	۳/۲۹۶	۳/۴۳۰	۴/۴۴۰	۳/۸۹۰	۴/۱۵۰
خوشه شماره ۴			۳/۸۵۵	۲/۶۲۶*	۳/۴۸۲	۳/۵۸۷	۴/۸۷۰	۲/۶۵۰*	۳/۵۱۵
خوشه شماره ۵				۴/۸۴۰	۴/۸۶۰	۴/۹۵۰	۵/۸۶۰	۲/۹۵۰	۲/۸۵۵*
خوشه شماره ۶					۲/۲۱۴*	۱/۲۱۸*	۳/۹۸۷	۳/۴۳۰	۴/۱۵۰
خوشه شماره ۷						۲/۵۱۵	۲/۴۱۵*	۳/۵۵۰	۳/۸۷۴
خوشه شماره ۸							۳/۸۷۰	۳/۹۰۵	۳/۹۸۷
خوشه شماره ۹								۴/۸۶۴	۴/۷۸۵
خوشه شماره ۱۰									۳/۹۸۰
کمترین مقدار هر ستون	۱/۰۲۲*	۱/۱۷۷*	۳/۴۸۲*	۲/۶۲۶*	۲/۲۱۴*	۱/۲۱۸*	۲/۴۱۵*	۲/۶۵۰*	۲/۸۵۵*

*کمترین مقدار هر ستون که بصورت زیرخط دار در جدول نشان داده شده است

** انتخاب بیشترین مقدار بین مقادیر زیر خط دار که بصورت تیره و زیر خط دار نشان داده شده است. این مقدار عدد خوشه بهینه است.

٤-۷ وزن‌دهی به شاخص‌های R,F,M

برای وزن‌دهی داده‌ها از خروجی نرم‌افزار اکسپرت چویس براساس مقایسه‌های زوجی شاخص بهره گرفته شد. طبق خروجی این نرم‌افزار، وزن‌های نسبی شاخص‌های ارزش پولی مبادلات (Monetary) معادل ۰/۴۹۳، تعداد تکرار مبادلات (Frequency) معادل ۰/۳۱۱ تازگی مبادلات (Regency) معادل ۰/۱۹۶ تعیین گردید (جدول ۴) که رابطه بین معیارهای اصلی در جدول ۵ نشان داده شده است.

جدول ٤. وزن پارامترهای RFM

پارامتر	تازگی مبادلات	تعداد تکرار مبادلات	ارزش پولی مبادلات
وزن	۰/۱۹۶	۰/۳۱۱	۰/۴۹۳

$$\begin{cases} W_T = W_{R+}\ W_{F+}\ W_{M-} \quad ۱ \\ W_T = ۰/۱۹۶ + ۰/۳۱۱ + ۰/۴۹۳ = ۱ \end{cases}$$

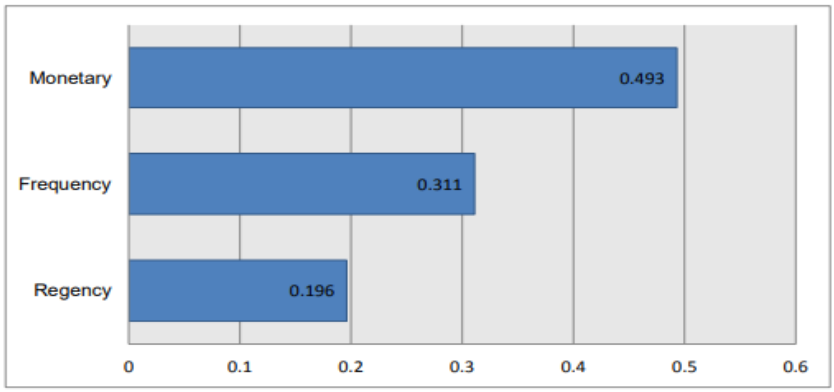
جدول۵. بررسی رابطه‌ای بین معیارهای اصلی

معیارهای اصلی	تازگی مبادلات	تعداد تکرار مبادلات	ارزش پولی مبادلات	وزن نهایی
تازگی مبادلات	۱	۲/۵۶	۲/۷۸	۰/۱۹۶
تعداد تکرار مبادلات		۱	۳/۸۹	۰/۳۱۱
ارزش پولی مبادلات			۱	۰/۴۹۳

بر این اساس میزان شاخص ناسازگاری ۰/۰۵ به‌دست آمد که نشان می‌دهد نتایج از روایی قابل قبولی برخوردار است.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]



شکل۵. نمودار وزن متغیرهای R,F,M با استفاده از روش AHP

در جدول ۶ مقادیر نرمال‌شده هر یک از پارامترهای RFM، ارزش طول عمر و رتبه خوشه‌ها بر اساس ارزش CLV آن‌ها مشخص شده است.

جدول ٦. محاسبه CLV و رتبه‌بندی مشتریان

خوشه	R نرمال	F نرمال	M نرمال	CLV	رتبه CLV	بخش (سگمنت)
C_1	۰/۹۹۶	۰/۰۲۵	۰/۰۰۲	۰/۱۴۱۴	۳	۳
C_2	۱	۰/۳۴۵	۰/۱۳۹	۰/۳۱۷۸	۱	۱
C_3	۰/۳۹۴	۰/۰۰۳	۰	۰/۰۵۳۳	۴	۴
C_4	۰/۹۹۹	۰/۱۵۹	۰/۰۱	۰/۱۸۸۰	۲	۲

با توجه به نتایج به‌دست آمده مشتریان در قالب هرم ارزش مشتری پیاده‌سازی شدند. خوشه شماره ۲ ارزش‌ترین خوشه در بین مشتریان است. سابقه ارتباط با این مشتریان طولانی مدت است و این تداوم ارتباط تاکنون نیز ادامه داشته است. این مشتریان کلیدی هستند و باید در حفظ و نگهداری آن‌ها تلاش شود تا ریزش ننمایند.

خوشه شماره ۴ مشتریانی هستند که نسبتا به تازگی و به صورت مداوم حجم قابل قبول اما نه خیلی زیاد مراجعات دارند.

خوشه شماره ۳ مشتریانی هستند که به تازگی به بانک مراجعه کرده‌اند و سابقه ارتباط طولانی مدت با آن‌ها وجود ندارد.

خوشه شماره ۱ مشتریانی هستند که دارای سابقه ارتباطات قبلی هستند ولی به تازگی ارتباط با آن‌ها کمرنگ‌تر شده است. این مشتریان در آستانه رویگردانی و ریزش هستند.

٤-٨ بهبود پیش‌بینی مشتریان رویگردان با استفاده از داده کاوی و سیستم خبره

عملکرد تکنیک‌های دسته‌بندی در پیش‌بینی رویگردانی را می‌توان با توجه به نرخ هشدارهای نادرست و نرخ فقدان هشدار مورد بررسی قرار داد. نرخ هشدار نادرست، نشان‌دهنده درصدی از غیر رویگردانان است که به نادرستی توسط تکنیک پیش‌بینی به عنوان رویگردان معرفی شده‌اند. نرخ فقدان هشدار، نشان‌دهنده درصدی از

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

پیش‌بینی‌های تکنیک مورد استفاده است که در شناسایی رویگردانان با شکست مواجه شده است. اگر نرخ هشدار نادرست بالا باشد، هزینه لازم برای نگهداری مشتریان وفادار افزایش می‌یابد؛ در صورتی که اگر نرخ فقدان هشدار بالا باشد، هزینه فرصت به علت عدم نگهداری و رویگردانی مشتریان افزایش می‌یابد[۱۹] و [۲۷] . با توجه به این که در این پژوهش، در سازمان مورد بررسی هزینه نگهداری یک مشتری بسیار کمتر از هزینه جذب مشتری جدید است، هدف روش مورد استفاده برای پیش‌بینی رویگردانی باید رسیدن به حداقل مقدار قابل دستیابی در نرخ فقدان هشدار باشد، درحالی که نرخ هشدار نادرست در سطح قابل قبولی نگه داشته شود. از این رو به منظور کاهش نرخ فقدان هشدار و افزایش کارایی سیستم پیش‌بینی، استفاده از یک سیستم ترکیبی براساس جنگل تصادفی، یکی دیگر از تکنیک‌های دسته‌بندی و نظر خبرگان پیشنهاد می‌شود.در چنین سیستمی یک مشتری زمانی به عنوان رویگردان معرفی می‌شود که هر دو تکنیک دسته بندی رأی به رویگردانی آن مشتری دهد و زمانی مشتری وفادار خواهد بود که رای هر دو تکنیک بر این مهم باشد. در چنین شرایطی تعداد مشتریان رویگردانی که مدل آن‌ها را به صورت نادرست وفادار معرفی کرده، کاهش می‌یابد. در شرایطی که یکی از مدل ها مشتری را به عنوان رویگردان و دیگری مشتری را به عنوان وفادار معرفی نماید، برچسب مشتری به نامشخص تغییر کرده و صدور رأی نهایی به عهده نظر خبرگان خواهد بود. در این صورت نرخ فقدان هشدار کاهش یافته و نرخ هشدار نادر ست در تعادل باقی می‌ماند. در مدل پیش‌بینی ترکیبی داده‌های مشتریان به صورت مستقل توسط هر کدام از تکنیک‌های جنگل تصادفی و درخت‌های فزاینده مورد بررسی قرار می‌گیرد.

هر مدل با توجه به ویژگی‌ها و پردازش داده‌های ورودی، یک مشتری به عنوان وفادار و یا رویگردان معرفی می‌نماید. به گونه‌ای که اگر مشتری در مدل مورد بررسی وفادار تشخیص داده شود، برچسب P_M =۱ به آن اختصاص داده می‌شود. در غیر این صورت برچسب مشتری P_M =۰ خواهد بود. پس از این‌که پردازش هر دو تکنیک و برچسب گذاری مشتریان پایان یافت، از یک سیستم ترکیبی برای انتخاب آرا استفاده می‌شود. در صورتی که هر دو مدل یک مشتری را رویگردان معرفی کرده باشند و $\begin{cases} P_{M_{Random\ forest}} = ۰ \\ \text{و} \\ P_{M_{Boosted\ Trees}} = ۰ \end{cases}$ در این صورت مدل ترکیبی به مشتری برچسب رویگردان و در صورتی که هر دو مدل مشتری را وفادار معرفی کرده و $\begin{cases} P_{M_{Random\ forest}} = ۱ \\ \text{و} \\ P_{M_{Boosted\ Trees}} = ۱ \end{cases}$ ، آنگاه مشتری را وفادار معرفی می‌کند. در صورتی که حداقل یکی از مدل‌های پیش‌بینی ، مشتری را به عنوان رویگردان شناسایی کند، مدل ترکیبی مشتری را در ریسک رویگردانی معرفی خواهد کرد. در شکل ۶ شمای مدل ترکیبی تشخیص ریسک رویگردانی نشان داده شده است.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

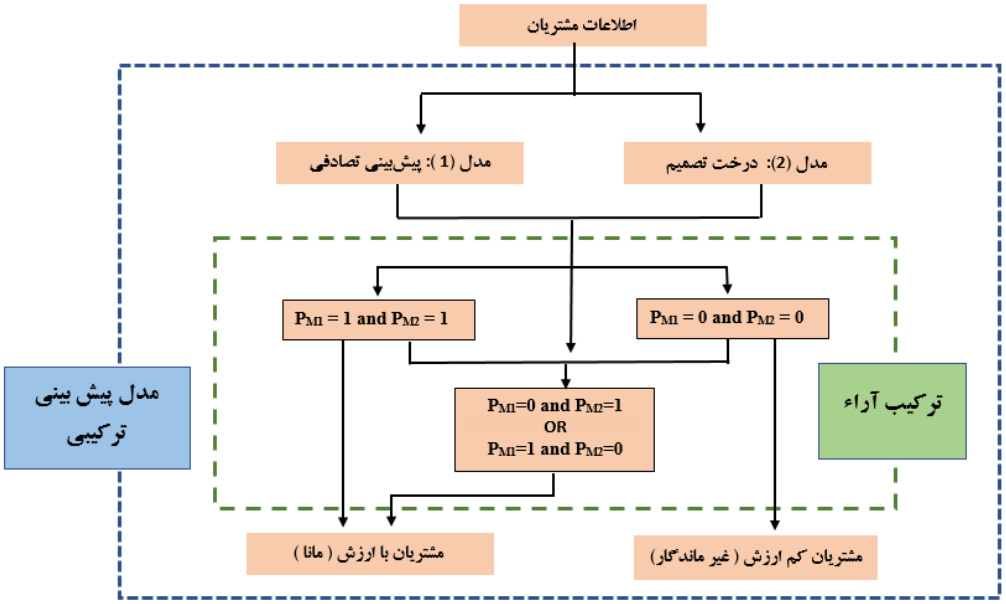
٤-٩ مقایسه مدل‌های معرفی‌شده

همان‌گونه که در بخش ۳-۱ اشاره شد، از سه مدل آموزش‌یافته به روش‌های جنگل‌های تصادفی، درخت‌های فزاینده و مدل ترکیبی با انتخاب منطقی جهت انجام بررسی ها استفاده شده است [۲۰-۱۸]صرف نظر از این نکته که دقت به‌دست آمده از این سه مدل تا حدودی می تواند نشان دهنده عملکرد بهتر هر مدل باشد، ولی نمی‌توان به صورت قطعی بر اساس دقت مدل‌ها با قطعیت بالا تصمیم گیری نمود. بنابراین به منظور افزایش قدرت تصمیم‌گیری و بررسی دقیق‌تر عملکرد هر مدل از روش زیر برای مقایسه بین مدل‌ها استفاده شده است.

رایج‌ترین ابزارهای مورد استفاده در دسته‌بندی مشتریان بر اساس رفتار رویگردانی، شبکه‌های عصبی و درخت تصمیم بوده‌اند. از بین راهکارهای پیاده‌سازی یک مدل دسته‌بندی مشتریان بر اساس رفتار آن‌ها، درخت تصمیم گزینه مورد نظر می‌باشد. یکی از مشکلات کاربرد درخت‌های تصمیم عدم ثبات و پایداری آن‌ها و ایجاد راه‌حل‌های بهینه محلی است. یکی از راهکارهایی که به منظور مقابله با مشکلات مذکور ارایه‌شده، استفاده از تکنیک‌های جدید دسته‌بندی جنگل تصادفی و درخت فزاینده می‌باشد. از نظر عملکرد پیش‌بینی این روش‌ها از بهترین تکنیک‌های دسته‌بندی‌اند.در روش جنگل تصادفی خروجی‌های دسته‌بندی کننده نسبت به داده‌های مغشوش و پرت بسیار توانمند است. با کاربرد این روش، اطلاعات مفیدی در زمینه میزان اهمیت هر یک از متغیرها به‌دست خواهد آمد و بدین ترتیب متغیرهایی که بیشترین تاثیر را بر متغیر وابسته دارند، تعیین می‌شوند.در این تکنیک مجموعه‌ای از درخت‌های تصمیم رشد داده می‌شود و هر درخت به مشهورترین کلاس رأی می‌دهد و با ادغام رأی درخت‌های مختلف، برای هر نمونه یک کلاس پیش‌بینی می‌شود. در این روش که برای افزایش صحت درخت تصمیم‌گیری طراحی شده است، تعداد بیشتری درخت تولید می‌شود تا برای پیش‌بینی کلاس با هم اقدام به رأی گیری کنند. در روش درخت‌های فزاینده نیز که با کاربرد روش افزایش دادن در درخت‌های تصمیم توسعه یافته است، به هر یک از نمونه‌های یادگیری وزنی اختصاص داده می‌شود و آموزش برای مجموعه‌ای ازK دسته‌بندی کننده تکرار می‌شود و به رأی هر دسته‌بندی‌کننده بر مبنای قدرت عملکر آن، وزنی اختصاص داده می شود. هر اندازه نرخ خطای یک دسته‌بندی کننده کمتر باشد، دقتش بیشتر خواهد بود و در نتیجه وزن بیشتری به رای آن تخصیص می یابد. کلاسی که مجموع وزنی بیشتری را به دست بیاورد، کلاس برنده خواهد بود و به عنوان کلاس پیش‌بینی شده معرفی می‌شود. با کاربرد این دقت پیش‌بینی افزایش می یابد. هنگام استفاده از کارکردهایی همچون دسته‌بندی که ماهیت پیش‌بینی دارند، ارزیابی مدل‌های استخراج شده ضروری است. برای این منظور، داده‌های مشتریان با ارزش به صورت تصادفی، با نسبت ۷۰ به ۳۰ به دو مجموعه داده‌های یادگیری و داده‌های ارزیابی تقسیم‌بندی شده و مدل حاصل از آن استنتاج می‌شود. سپس دقت مدل با استفاده از مجموعه داده‌های ارزیابی اندازه‌گیری می‌شود. به منظور بررسی دقت عملکرد و اثر بخشی مدل پیش‌بینی، در این مطالعه از روش ماتریس آشفتگی و منحنی تجمعی منفعت استفاده شده است. در انتها برای افزایش اثربخشی و دقت مدل‌های مورد استفاده، از یک مدل ترکیبی نیز استفاده شده است که با استفاده از ترکیب رای‌هایی که از هر دو مدل به‌دست آمده است، به انتخاب کلاس برنده می‌پردازد.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]



شکل٦. مدل ترکیبی تشخیص ریسک رویگردانی مشتریان با ارزش با استفاده از داده کاوی و سیستم خبره

٤-١٠ تعیین مجموعه داده‌های آزمون و توزیع مربع کای

به منظور افزایش قدرت تصمیم‌گیری و بررسی دقت در عملکرد هر مدل، از روش آزمون فرض‌های آماری برای مقایسه بین مدل‌ها استفاده شده است. یکی از مهم‌ترین مسایل موجود در آمار استنباطی، آزمون فرض‌های آماری است که در تصمیم‌گیری بسیار حایز اهمیت است. یکی از مسایل موجود در آزمون فرض ارایه یک معیار برای پذیرفتن و یا رد کردن فرض است. یکی از معیارهایی که جهت انجام آزمون فرض معرفی شده است، ارزش P (P-Value) است. یک فرض آماری، ادعایی در مورد یک یا چند جمعیت مورد بررسی است که ممکن است درست یا نادرست باشد. به عبارت دیگر فرض آماری یک ادعا یا گزاره‌ای در مورد توزیع یک جمعیت یا پارامتر توزیع یک متغیر تصادفی است. برای بررسی یک فرض آماری پس از انجام یک آزمایش تصادفی، اگر نتیجه آزمایش تفاوت «معنی داری» با آنچه که مورد انتظار است، داشته باشد در صورتی که فرض مورد نظر را صحیح فرض شده باشد، فرض را رد کرده و در غیر این صورت پذیرفته می‌شود. بنابراین با طرح یک آزمون آماری در این جا برای هر سه مدل و محاسبه ارزش P و با تفسیر آن به مقایسه بین مدل‌ها پرداخته می‌شود [۱۹].

از آنجا که هیچ پژوهش دیگری بر روی داده های مورد استفاده پژوهش حاضر اجرا نشده و نتایج قابل مقایسه‌ای وجود ندارد، معیار بررسی های انجام‌شده ، نتایج مقادیر مربع کای و ارزش P برای سه الگوریتم مورد بحث این پژوهش خواهد بود.

به منظور تعیین داده‌های مورد نیاز برای این آزمون آماری، مجموعه داده‌های مورد استفاده بر اساس اعتبارسنجی متقاطع nتایی به طور تصادفی به مجموعه‌های داده محدودی با اندازه مساوی شکسته می‌شود. هر یک از این مجموعه ها در ادامه به عنوان یک مجموعه اعتبارسنجی به کار خواهد رفت. از ۱۵۸۴ داده مورد استفاده در این مطالعه ۵۲ مجموعه داده با اندازه ۳۰ برای مجموعه آموزش و آزمایش ایجاد شده است. دقت هر

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]
[DOI: 10.71773/jamlu-2025-1-2127]

مدل برای هم مجموعه داده‌های آزمون و هم مجموعه داده‌های آموزش به صورت جداگانه محاسبه شده است. بر اساس دقت به دست آمده از مدل در مجموعه داده‌ها با استفاده از توزیع مربع کای بر اساس رابطه زیر، مقدار ارزش P برای هر مدل اندازه‌گیری شده است.

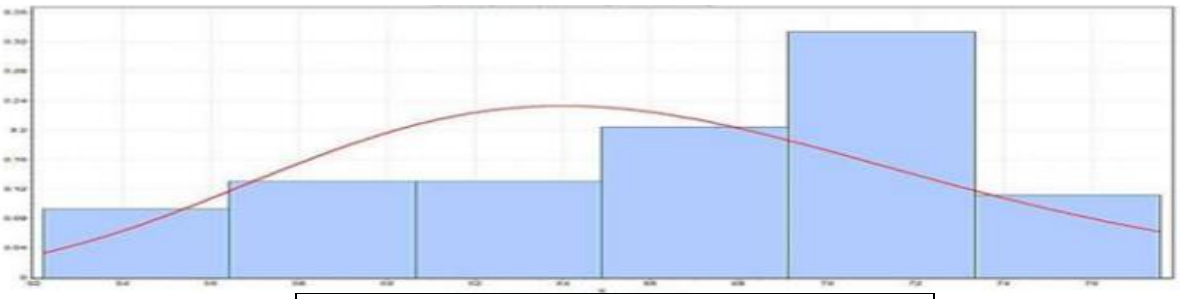
$$X^2 = \sum_{All\ Data} \frac{(Observed - Expected)^2}{Expected} \qquad (10)$$

مقادیر X^2 برای هر سه مدل برابر جدول ۷ می‌باشد.

جدول ۷. نتایج مقادیر مربع کای و ارزش P برای اجرای الگوریتم‌ها

ارزش P	مربع کای	الگوریتمهای مورد استفاده
P-value $_{RF}$ ≈./۲	$X^2_{Random\text{-}Forest}$(r=5) =۸۵۶۴۳	جنگل تصادفی
P-value $_{BT}$ ≈ ./۲	$X^2_{Boosted\ Tree}$ (r=5) = ۶۶۴۵۲	درخت فزاینده
P-value $_{Hybrid}$ ≈ ./.۲	X^2_{Hybrid} (r=3) = ۹۵۷۷۵	ترکیبی

بر اساس مقادیر ارزش P به دست آمده برای هر مدل، مشخص شد که مدل ترکیبی برای پیش‌بینی مدل رویگردانی مشتریان بهتر عمل می‌کند. در شکل‌های ۷و ۸ و ۹ نمودار توزیع داده‌ها توسط مربع کای نیز نشان داده شده است.

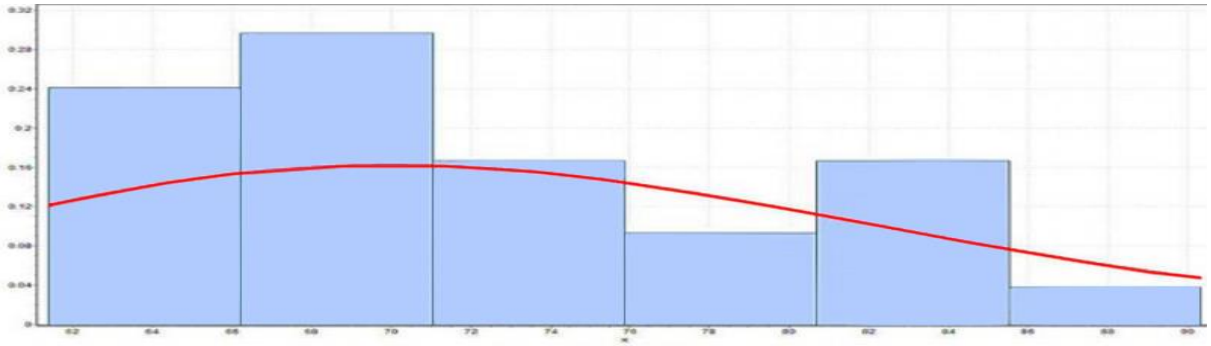


مربع کای					
۵				درجه آزادی(Deg.Freedom)	
۸/۵۶۴۳				ساکن (static)	
۰/۲۶۹۱۳				ارزش p	
۰/۰۱	۰/۰۲	۰/۰۵	۰/۱	۰/۲	
۱۳/۲۷۷	۱۱/۶۶۸	۹/۴۸۷۷	۷/۷۷۹۴	۵/۹۸۸۶	مقدار بحرانی

شکل ۷. نمودار توزیع مربع کای برای جنگل تصادفی

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

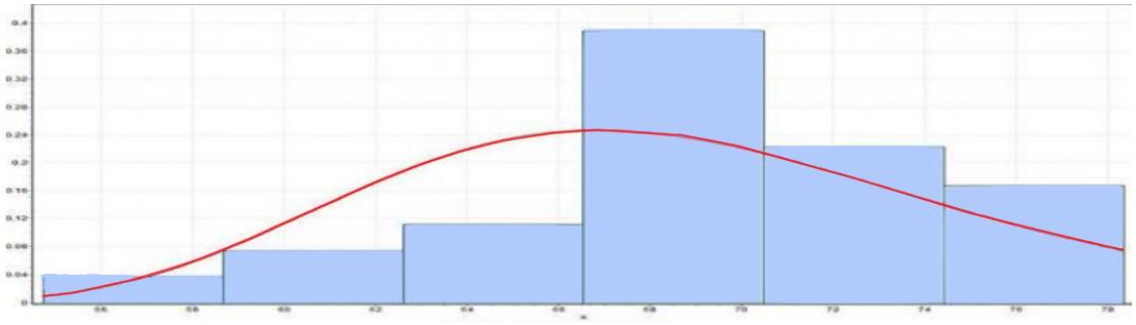
[DOI: 10.71773/jamlu-2025-1-2127]



مربع کای					
۳				درجه آزادی	
۹/۵۷۷۵				ساکن	
۰/۰۲۲۵				ارزش p	
۰/۰۱	۰/۰۲	۰/۰۵	۰/۱	۰/۲	
۱۱/۳۴۵	۹/۸۳۷۴	۴/۸۱۴۷	۶/۲۵۱۴	۴/۶۴۱۶	مقدار بحرانی

شکل ۸. نمودار توزیع مربع کای برای درخت‌های فزاینده

تفکیک مشتریان رویگردان از وفادار را می‌توان با کاربرد هر یک از روش‌های دسته‌بندی انجام داد. از آنجایی که هدف نهایی نگاشت مشتریان به دو گروه مانا و رویگردان خواهد بود، نیاز به یکی از مدل‌های پیش-گویانه می‌باشد. داده کاوی پیش‌گویانه مدلی از سیستم ارائه می‌دهد که شامل به کارگیری متغیرها و فیلدها در انبار داده جهت پیش‌گویی مقادیر ناشناخته می‌باشد.



مربع کای					
۵				درجه آزادی	
۶/۶۴۲۵				ساکن	
۰/۲۴۸۶۲				ارزش p	
۰/۰۱	۰/۰۲	۰/۰۵	۰/۱	۰/۲	
۱۵/۰۸۶	۱۳/۳۸۸	۱۱/۰۷	۹/۲۳۶۴	۷/۲۸۹۳	مقدار بحرانی

شکل ۹. نمودار توزیع مربع کای برای مدل ترکیبی

در این مطالعه به منظور شناسایی مشتریان رویگردان و وفادار در سازمان، تمرکز اصلی بر گروه مشتریان با ارزش می‌باشد. ارزش گذاری مشتریان در سازمان بر اساس مدل RFM صورت می‌پذیرد که در این بخش‌بندی، ارزش‌گذاری مشتریان بر اساس مشخصه‌های تاخر(مدت زمان سپری شده در آخرین تراکنش مشتری با سازمان)،

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

تناوب خرید (تعداد تراکنش‌های مشتری در بازه زمانی مشخص) و ارزش پولی (میزان پول خرج شده توسط مشتری در سازمان در یک بازه) محاسبه می‌شود. مشخصه‌های RFM ، پیش‌بینی کنندگان خوبی برای ارزش عمر مشتریان هستند. این متغیرها به روش‌های متفاوتی برای ارزش‌گذاری مشتریان به کار گرفته می‌شوند. در یکی از این روش‌ها، هر کدام از متغیرهای R، F و M دارای ارزش یکسانی هستند. در یک رویکرد جدید در این مدل، به هر یک از متغیرها وزنی اختصاص داده می‌شود. سپس بر اساس وزن‌های مورد نظر ارزش‌گذاری مشتریان انجام می‌شود. از آنجایی که تعیین اوزان بر اساس نظر خبرگان سازمان صورت می‌پذیرد به آن رویکرد RFM قضاوت محور گفته می‌شود که در این مطالعه از این رویکرد استفاده شده است.

۵ نتیجه‌گیری

در این پژوهش به دلیل اهمیت وفاداری مشتری و همچنین ضرورت تشخیص احتمال رویگردانی مشتری برای محاسبه‌ی ارزش دوره عمر وی، به بررسی و مدیریت رویگردانی مشتریان پرداخته‌شده و پیاده‌سازی مدل رویگردانی مشتریان و نیز نحوه طراحی یک سیستم خبره در صنعت خودرو ارایه شده است. بررسی نتایج پژوهش‌های پیشین حاکی از آن است که طول ارتباط فعال با مشتری، متوسط فاصله زمانی بین خرید مشتریان و تناوب نسبی، بهترین پیش‌بینی کنندگان برای جداسازی مشتریان وفادار از مشتریان متمایل به رویگردانی می‌باشند. در این مطالعه، با بهره‌گیری از تکنیک دسته‌بندی در داده کاوی بر پایه مدل RFM و طبق روش‌شناسی استاندارد CRISP-DM، با بررسی پایگاه داده یکی از شرکت‌های خودروساز، به طراحی مدلی برای پیش‌بینی مشتریان رویگردان در سازمان پرداخته شد. با تمرکز بر گروه مشتریان با ارزش، ارزش‌گذاری مشتریان در سازمان بر اساس مدل قضاوت محور RFM مبتنی بر نظر خبرگان صورت پذیرفت. برای دسته‌بندی مشتریان بر اساس رفتار رویگردانی، از تکنیک‌های دسته‌بندی جنگل تصادفی و درخت فزاینده و یک مدل ترکیبی(ترکیب دسته‌بندهای جنگل تصادفی، درخت فزاینده و نظر خبرگان) استفاده شد و مقادیر مربع کای و P-value برای سه الگوریتم مورد بحث این پژوهش بررسی شد. مقادیر P-value برای دو مدل جنگل تصادفی و درخت فزاینده حدود مقدار ۰/۲ و برای مدل ترکیبی پیشنهادی ۰/۰۲ به‌دست آمد که نشان‌دهنده عملکرد بهتر مدل ترکیبی پیشنهادی برای پیش‌بینی مدل رویگردانی مشتریان است.

منابع

[1] Hill, C.W. L. and Jones, G.R., (2008). Strategic management theory: An integrated approach, Houghton Mifflin Company, Boston, New York.

[2] Sahay, A., (2013). A customer oriented approach to identifying competitive advantage. Indian Institute of Management, 5(8).

[3] Shamsuzzaman, M.D., (2013). Customer satisfaction case: National life insurance company, Laurea University of Applied Sciences.

[4] Salehabadi, A., Najafi, J. , Sayar, M., Salari, M., (2014). Economic evaluation of several simultaneous projects, considering their effects on each other, International Conference on New Researches in Management, Economics and Accounting, 29 Azar 2014 (in Persian)

[5] Donkers, B., Verhoef, P.C. and de Jong, M.G., (2007). Modeling CLV: A test of competing models in the insurance industry. Quantitative Marketing and Economics, 5(2), 163-190.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]

[DOI: 10.71773/jamlu-2025-1-2127]

[6] Jain, D. and Singh, S.S., (2002). Customer lifetime value research in marketing: A review and future directions. Journal of interactive marketing, 16(2), 34-46.
[7] Glady, N., Baesens, B. and Croux, C., 2009. Modeling churn using customer lifetime value. European Journal of Operational Research, 197(1), 402-411.
[8] Nie, G., (2011). Credit card churn forecasting by logistic regression and decision tree. Expert Systems with Applications 3812, 15273-15285.
[9] Xie, Y., (2009). Customer churn prediction using improved balanced random forests. Expert Systems with Applications 363, 5445-5449.
[10] Lin, Ch., Gwo-Hshiung T., and Yang-Chieh Ch. (2011). Combined rough set theory and flow network graph to predict customer churn in credit card accounts. Expert Systems with Applications 381, 8- 15.
[11] Grieger, M., Ludwig, A., (2019). On the move towards customer-centric business models in the automotive industry – a conceptual reference framework of shared automotive service systems. Electron Markets 29, 473-500.
[12] Balinado, Justine R., Yogi T. Prasetyo, Michael N. Young, Satria F. Persada, Bobby A. Miraja, and Anak A.N. Perwira Redi. (2021). The Effect of Service Quality on Customer Satisfaction in an Automotive After-Sales Service, Journal of Open Innovation: Technology, Market, and Complexity 7(2), 116.
[13] Nassiri Pirazari, k., Jalilian, k. (2020), Designing an Optimal Customer Satisfation Model in Automotive Industry, J Control Autom Electr Syst 31, 1-39.
[14] Sharaj Shafiei, A, (2011). Investigating the relationship between after-sales services of automotive companies (Saipa and Iran Khodro) and customer satisfaction, Journal of Industrial Strategic Management, 8, 9-18, Spring 2011(in Persian)
[15] Edi, L., . Krajnović, A. and Jurica Bosna, J., (2022). Sustainability Strategies and Achieving SDGs: A Comparative Analysis of Leading Companies in the Automotive Industry, Sustainability 14, no. 7: 4000. https://doi.org/10.3390/su14074000
[16] Juan,L., José Moral, M. , (2020). Automotive Aftermarket Forecast in a Changing World: The Stakeholders' Perceptions Boost!, Sustainability 12, no. 18: 7817. https://doi.org/10.3390/su12187817
[17] Ebrahimi,A., Bakhshizadeh, P., Varasteh, R., (2022). A predictive analytics approach to improve the dealers-manufacturer relationship in the after-sales service network; case study in the automotive industry, International Journal of Management Science and Engineering Management.
[18] Baines, P., Fill, Ch., & Page, K., (2010). Market segmentation and posintioning. Marketing, (Secon d Edi). USA: Oxford University Press: 227 - 244.
[19] Hiziroglu, A., Sisci, M., Cebeci, Halil,I., cebeci . , S., (2018). An empirical assessment of customer li fetime value models within data mining. Baltic J. Modern Computing, 6 (4), 434-
448 https://doi.org/10.22364/bjmc.2018.6.4.08 .
[20] Hughes, A.M. (2012). Strategic database marketing 4 e: the masterplan for starting and managing a profitable, customer - based marketing program, Publisher(s): McGraw-Hill,ISBN:9780071773645.
[21] Jiawei, H., Micheline, K., (2006). Data mining: concepts and technigues. Annals of Physics, 54,770.
[22] Berry, J., Linoff, G.S., (2004), Data mining techniques for marketing, sales and customer relationship management". John Wiley & Sons, 2nd edition.
[23] Gupta, S., Lehmann, D.R. and Stuart, J.A., (2004). Valuing customers. Journal of marketing research, 41(1), 7-18.
[24] Sellers, J. and Hughes, A., (1994). RFM Analysis: A New Approach to a Proven Technique. www. relation-shipmktg. com/Free Articles/rmr017.pdf.
[25] Hadden, J., (2005), Computer assisted customer churn management: State-of-the-art and future trends. Computer & Operations Research, 34, 2902–2917.
[26] Witten, I.H., Frank, E., Hall, M.A., Pal, C.J. and DATA, M., (2005). Practical machine learning tools and techniques. In DATA MINING, 2, 4.
[27] Yufang, L., Deheng, Sh., (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value, 42, 387-400. https://doi.org/10.1016/j.im.2004.01.008.

[Downloaded from jamlu.lahijan.iau.ir on 2026-05-13]
[DOI: 10.71773/jamlu-2025-1-2127]