

طراحی مدل تخصیص خدمات بانکی

تاریخ دریافت: ۱۴۰۳/۰۲/۱۲

تاریخ پذیرش: ۱۴۰۳/۰۹/۲۴

نازنین کاشانی نیکو*

مهناز ربیعی**

کیامرث فتحی هفشجانی***

چکیده

امروزه نحوه تعامل سازمان ها و بخصوص بانک ها با مشتریان در قالب مدیریت ارتباط با مشتری (CRM) به طور قابل توجهی تغییر یافته است. بر این اساس هدف تحقیق حاضر تبیین مدل ارتباط با مشتریان با رویکرد سیستم های بیزینس-توصیه گر فازی است. تحقیق حاضر از نظر مخاطب، کاربردی، از نظر هدف، اکتشافی و از نظر زمانی در دسته تحقیقات مقطعی می باشد. در این مطالعه اطلاعات ۹۸۶۰۴ مشتری بررسی شده است. در این تحقیق از ۳ رویکرد داده کاوی، فازی، میانگین گیری بیزین غیر خطی بهره گرفته شده است. در این تحقیق اطلاعات ۲۲ شاخص مرتبط با مشتریان وارد مدل های بیزین غیر خطی (BMA, TVP-DMA و TVP-DMS)، گردید. بر اساس میزان خطا، مدل BMA از بالاترین دقت برخوردار بود. بر اساس نتایج بدست آمده، متغیرهای موجودی حساب؛ مجموع مانده سپرده ها؛ مجموع مانده کل تسهیلات جاری و حجم تراکنش های مالی به عنوان متغیرهای غیر شکننده شناسایی شدند. در ادامه بر اساس نتایج، رویکرد C-MEANS نسبت به K-MEANS از دقت بالاتری برخوردار است. بر اساس رویکرد C-MEANS، ۱۶ خوشه شناسایی گردیدند و ویژگی های هر ۱۶ خوشه مورد تجزیه و تحلیل قرار گرفت. بر اساس نتایج اکثر متغیرها تأثیر مثبت و معناداری بر وضعیت هر خوشه دارند.

واژگان کلیدی: خدمات بانکی، مدیریت ارتباط با مشتریان، خوشه بندی فازی، داده کاوی.

* دانشجوی دکتری، گروه مدیریت فناوری اطلاعات، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران
** گروه اقتصاد، مرکز تحقیقات مدل سازی و بهینه سازی علوم مهندسی، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران، (Mahnaz.rabiei@yahoo.com)
*** گروه مدیریت صنعتی، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران

مقدمه

صنعت بانکداری به دلیل سیستم‌های اطلاعاتی و فناوری‌های نوآورانه‌اش شناخته شده است (کینگی^۱ و همکاران، ۲۰۲۲). ظهور اینترنت به بانک‌ها کمک کرده تا حجم وسیعی از داده‌ها را جمع‌آوری و استفاده کنند (ویدال^۲ و همکاران، ۲۰۲۰). مدیران بانک‌ها به تدریج متوجه شدند که اطلاعات و داده‌ها یکسان نیستند (تئودوریس^۳ و همکاران، ۲۰۲۲). داده‌ها قبل از تبدیل شدن به اطلاعات معنی دار و قابل استفاده در تصمیم‌گیری باید تجزیه و تحلیل شوند (زین‌العابدین^۴ و همکاران، ۲۰۲۳). یکی از مهم‌ترین مدل‌های طبقه‌بندی مشتریان بانک، مهندسی ویژگی است که فرآیند تبدیل داده‌های خام به ویژگی‌های جدید و انتخاب بهترین ویژگی‌ها است (لیو و همکاران^۵ ۲۰۲۲؛ یوان و همکاران^۶ ۲۰۲۲؛ اسلام و همکاران^۷ ۲۰۲۲).

اخیراً، تکنیک‌های مختلف دانش کاوی و داده کاوی برای کشف اطلاعات پنهان در مورد رفتار مشتری مورد استفاده قرار گرفته‌اند (کالایوانی و سوماتی^۸ ۲۰۱۹؛ دی کاپینی و همکاران^۹ ۲۰۲۰؛ جین و همکاران^{۱۰} ۲۰۲۱؛ چن و همکاران^{۱۱} ۲۰۲۱؛ آلام و همکاران^{۱۲} ۲۰۲۱). کشف دانش به تصمیم‌گیری در بسیاری از زمینه‌های تحقیقاتی کمک کرده است. دانش کاوی می‌تواند خدمات بانکی را برای مشتریان قابل قبول‌تر و قابل اعتمادتر نماید (کلرکین و هانسون^{۱۳} ۲۰۲۱؛ برگرون و همکاران^{۱۴} ۲۰۲۰).

روش‌های مختلف طبقه‌بندی مبتنی بر یادگیری ماشین برای استخراج دانش در مورد مشتریان بانک‌ها استفاده می‌شوند؛ تا بتوانند مشتریان ارزشمند را شناسایی کرده و از ریزش آن‌ها جلوگیری کنند. امین و همکاران^{۱۵} (۲۰۱۹)، از روش (NB^{۱۶}) برای پیش‌بینی ریزش مشتری در حالی که عدم قطعیت داده‌ها را در نظر گرفت، استفاده کرد. روش‌های یادگیری ماشینی نتایج امیدوارکننده‌تری در این حوزه ارائه نموده‌اند (لیو و همکاران، ۲۰۲۲). انتخاب ویژگی یک مرحله پیش پردازش حیاتی است که به طبقه‌بندی مشتریان بانک می‌نماید (یوان و همکاران ۲۰۲۲؛ اسلام و همکاران ۲۰۲۲). مهندسی ویژگی به طور قابل توجهی بر دقت پیش‌بینی روش‌های طبقه‌بندی تأثیرگذار است (بنسان و همکاران^{۱۷} ۲۰۱۶؛ لانگ و همکاران^{۱۸} ۲۰۱۹؛ عابدین^{۱۹} و همکاران ۲۰۲۰؛ ژانگ و همکاران^{۲۰} ۲۰۲۱ الف؛ ژانگ و همکاران ۲۰۲۱ ب)؛ بنابراین، هدف اصلی این مقاله پیش‌بینی رفتار و طبقه‌بندی مشتریان بانک بر اساس این رفتارها است.

برخلاف مطالعات موجود که مرحله پیش پردازش داده خود را به انتخاب ویژگی محدود نمودند (کرامتی و همکاران^{۲۱} ۲۰۱۶؛ یوان و همکاران ۲۰۲۲)؛ این مطالعه روش‌های تبدیل ویژگی را نیز در نظر می‌گیرد.

پیشینه پژوهش

در حال حاضر حجم زیادی از داده‌ها در بانک‌ها ذخیره شده است و به طور مداوم و با سرعت زیادی در حال رشد است. مانع اصلی تبدیل این حجم عظیم از داده‌ها به اطلاعات آموخته شده و پایگاه داده و ایجاد یک هوش رقابتی است. چندین مطالعه از داده‌های مشتریان بانک برای شناسایی الگوهای رفتاری با استفاده از تکنیک‌های مختلف دانش کاوی و داده کاوی استفاده کرده‌اند.

بانک‌ها دارای تعداد زیادی بانک‌های اطلاعاتی عظیم برای تراکنش‌ها و سایر جزئیات مرتبط با مشتریان خود است. میتوان از این داده‌های ذخیره شده اطلاعات کسب و کار با ارزشی را استخراج کرد؛ اما از این داده‌های خام برای پشتیبانی از تجزیه و تحلیل و تصمیم‌گیری با استفاده از زبان پرس و جوی سنتی استفاده نمی‌شود چون تجزیه و تحلیل انسان با این حجم و بعد گسترده منجر به شکست می‌شود. روش‌های آماری سنتی ظرفیت و مقیاس برای تحلیل این حجم داده نداشتند و از این رو استفاده روش‌های مدرن داده کاوی و ابزارهای آن به طور گسترده ای رو به افزایش است. این روش‌ها نه تنها در بانک‌ها و مؤسسات مالی بلکه در کل صنایع نیز نفوذ پیدا کرده‌اند. مزیت داده کاوی این است که می‌تواند حجم زیادی از اطلاعات را مدیریت کند و ساختارهای ذاتی الگوها در داده را آموزش بدهد داده کاوی می‌تواند قوانین و مدل‌هایی که در ساختار تصمیم‌گیری برای موردهای آینده استفاده می‌شود را تولید کند مدل سازی رفتار با مشتری و یا پروفایل مشتری یک ابزار برای پیش بینی ارزش آینده از یک فرد خاص به گروه ریسک بر اساس ویژگی های جمعیت شناختی، خود سبک زندگی و رفتار قبلیش است که به حفظ مشتری کمک می‌کند. تصمیم‌گیری صحیح و ایجاد استراتژی‌ها و تاکتیک‌های مؤثر بازاریابی و فروش، منوط به شناسایی و درک صحیح سازمان‌ها از مصرف‌کنندگان و خریداران است. محققان بازاریابی به طور دائم در تلاشند تا رفتار مصرف‌کنندگان را شناسایی کنند و بر این اساس، راه کارهایی مناسب را

برای فروش بهتر و مؤثرتر و افزایش سهم بازار اتخاذ کنند. یکی از مهم ترین فعالیت های بانک ها شناسایی مشتریان و ارائه خدمات متنوع بانکی متناسب با ویژگی های آنها است تا بتوانند به افزایش سهم خود در بازار دست یابند. تحلیل رفتار مشتریان و ارتباط مناسب با آنها می تواند عامل مهمی برای بقای بانک ها در عرصه رقابت تلقی شود. هدف از شناسایی مشتریان ایجاد تمایز بین آنها و تشخیص مشتریان پرازش نگهداری آنها و جذب مشتریان ارزشمند است.

هنگام انتخاب الگوریتم ها برای روش های طبقه بندی، رایج ترین رویکرد مرسوم شامل دانش متخصصان، روش آزمون و خطا یا تحلیل نظری موضوعات مورد بررسی است. به گفته وانگ و همکاران (۲۰۱۵)، چنین پیشنهادهایی دارای معایب زیر است: هزینه های محاسباتی بالا در مورد مجموعه داده های بسیار بزرگ، ناتوانی در به دست آوردن دانش در مورد همه طبقه بندی کننده های ناشی از ارزیابی خطاهای نمایشی آنها و با وجود امکان همکاری با کارشناسان میدانی، مانند یک راه حل همچنین مستلزم روابط مالی و صحیح قابل توجه با متخصصان است. خان و همکاران (۲۰۲۰)، نشان دادند که افزایش قابل توجهی در تقاضا برای سیستم های یادگیری ماشین وجود دارد که می تواند فرآیند انتخاب الگوریتم های مناسب را با توصیه آنها برای کارهای مختلف خودکار کند. به نظر آنها، چنین سیستم هایی معایب رویکردهای مرسوم را ندارند و امکان استفاده از الگوریتم های یادگیری ماشین را برای حل مشکلات جدید فراهم می کنند و همچنین به افراد غیر متخصص اجازه می دهند تا به طور مستقل در این حوزه دست به پژوهش بزنند. فرآیند شناسایی ویژگی مشتریان و طبقه و دسته بندی مشتریان از یک تاریخچه طولانی برخوردار است. در ادامه به برخی از تحقیقات خارجی و داخلی صورت گرفته در این حوزه اشاره شده است.

موجیکا و همکاران^{۲۲} (۲۰۰۲)، یک سیستم را برای تجزیه و تحلیل داده ها در مورد رفتار مشتریان بانک پیشنهاد کردند. در سیستم طراحی شده، موارد با استفاده از نقشه های ویژگی خودسازمان دهی که به صورت بدون نظارت آموزش داده شدند استخراج می شد و این امر منجر به یک سیستم ساده و کارآمد شد. با این حال، این رویکرد به دلیل دشواری تطبیق موارد موجود در پایگاه داده با مشتریان جدید؛ چندان مورد توجه واقع نشد. وینارسکی^{۲۳} (۲۰۰۲)، همان مجموعه داده را با استفاده از یک طبقه بندی تابع انتقال محلی

برای شناسایی مشتریان فعال و غیرفعال بانک مورد بررسی قرار داد. اطلاعات دریافتی در لایه پنهان معماری شبکه عصبی تحت نظارت برای استخراج الگوهای رفتاری از پایگاه داده استفاده شدند. با این حال، طبق گزارش‌ها، این مدل مستعد برازش بیش از حد بود. باومن و همکاران^{۲۴} (۲۰۰۷)، از سوابق ۱۹۵۱ مشتری بانک استفاده کردند و مدل‌های مختلفی را برای پیش‌بینی سهم کیف پول^{۲۵} (SOW) برای سپرده‌ها، بدهی‌ها و وام‌ها طراحی نمودند. با این حال، این مدل‌ها فقط بر اساس رگرسیون لجستیک بودند، که اجازه استخراج روابط غیرخطی بین ویژگی‌های ورودی و خروجی مشتریان بانک را نمی‌داد. نگای و همکاران^{۲۶} (۲۰۰۹)، طبقه‌بندی مطالعات قبلی را بر اساس روش‌های داده‌کاوی برای وظایف مختلف مدیریت ارتباط با مشتری بررسی کردند. در این رویکرد از روش‌های خوشه‌بندی برای شناسایی مشتریان و از روش‌های طبقه‌بندی برای جذب و حفظ مشتریان استفاده شد. کرامتی و همکاران (۲۰۱۶)، مجموعه داده انحراف مشتری را با استفاده از تکنیک‌های داده‌کاوی مورد بررسی قرار دادند. از یک درخت تصمیم برای شناسایی ریزش مشتری در خدمات بانکداری الکترونیکی استفاده نمودند. روش انتخاب ویژگی به عقب برای حذف ویژگی‌های ورودی نامرتبب استفاده شد. فیجزا و همکاران^{۲۷} (۲۰۱۷)، مدلی از رفتار مصرف‌کننده در بانکداری را پیشنهاد کردند که تئوری رفتار مصرف‌کننده را توضیح می‌داد. یک مدل یادگیری ماشینی توسط رحمان و خان^{۲۸} (۲۰۱۸)، برای تشخیص رفتار مشتری با تجزیه و تحلیل مجموعه داده بازاریابی بانکی دانشگاه کالیفرنیا ارواین^{۲۹} (UCI) پیشنهاد شد. کالایوانی و سوماتی^{۳۰} (۲۰۱۹)، میلیون مشتری خروجی را با استفاده از روش‌های DT^{۳۱}، MLP^{۳۲}، SVM، NB و LR^{۳۳} مورد بررسی قرار دادند تا گام‌های استراتژیک برای جلوگیری از ریزش مشتری بردارند. برای کاهش ابعاد داده‌ها، از تحلیل عاملی و تحلیل مؤلفه‌های اصلی به عنوان تکنیک‌های استخراج ویژگی سنتی استفاده شد. با این حال، در مقایسه با تکنیک‌های انتخاب ویژگی، رویکرد آن‌ها به از دست دادن اطلاعات بالقوه مرتبط و کاهش قابلیت تفسیر ویژگی‌ها منجر شد. برای شناسایی مشتریان ارزشمند بانک، ژو و همکاران^{۳۴} (۲۰۱۹)، یک مدل یادگیری ماشینی را پیشنهاد کردند که در آن از روش‌های استخراج قانون MLP و انجمن برای تجزیه و تحلیل رفتار مشتریان بانک استفاده شد. روش مبتنی بر شبکه عصبی به دلیل ظرفیت آن در مدل‌سازی روابط رفتاری غیرخطی در داده‌ها، از نظر طبقه‌بندی دقیق‌تر بود. راجو و داندایودام^{۳۵} (۲۰۱۸)، همچنین مجموعه داده‌های بانک

برچسب‌گذاری شده را با استفاده از روش‌های NB، DT و MLP برای بررسی الگوهای رفتاری برای مدیریت موثر ارتباط با مشتری تجزیه و تحلیل کردند. در این مورد، درخت تصمیم بهترین عملکرد را داشت و این را می‌توان به توانایی آن در برخورد موثر با متغیرهای طبقه‌بندی نسبت داد. عباسی مهر و شعبانی^{۳۶} (۲۰۱۹)، از خوشه‌بندی سری‌های زمانی برای شناسایی چهار دسته از مشتریان کسب و کار بانکی به عنوان مشتریان با ارزش بالا، مشتریان با ارزش متوسط، مشتریان مستعد ریزش و ریزش کنندگان استفاده کردند. سود اضافی به دست آمد و اطلاعات بالقوه مرتبط از اسناد متنی مشتری استخراج شد تا یک کمپین حفظ مشتری موفق آماده شود (دی کاینی و همکاران^{۳۷}، ۲۰۲۰). هو و همکاران^{۳۸} (۲۰۱۹)، یک مدل یادگیری ماشینی از رفتار مشتری برای اطمینان از امنیت داده برای مشتریان بانک پیشنهاد کردند و این فرصت‌های بیشتری را برای استفاده از داده‌ها در مورد رفتار مشتری بانک پیشنهاد کرد. یک سیستم مدیریت ارتباط با مشتری توسط چن و همکاران^{۳۹} (۲۰۲۱)، توسعه داده شد. عمران خان و همکاران (۲۰۲۳)؛ به بررسی رفتار و فعالیت مشتری در بخش بانکداری با استفاده از اطلاعات ۲۴۰۰۰ مشتری فعال و غیر فعال پرداختند و از تکنیک‌های مختلف تبدیل ویژگی برای تبدیل داده‌های رفتاری استفاده نمودند. سپس در ادامه با انتخاب ویژگی‌ها اقدام به طبقه‌بندی مشتریان نمودند و پس از مدل‌سازی رفتار مشتریان بانک‌ها بسته به طبقه هر مشتری سیاست‌هایی برای افزایش فعالیت مشتریان و افزایش رضایت‌مندی آن‌ها پیشنهاد نمودند.

برای خلاصه کردن ادبیات فوق، تحقیقات تا به امروز بیش‌تر بر روی روش‌های پیش‌بینی مبتنی بر یادگیری ماشین متمرکز بوده‌اند تا اینکه اقدام به هوشمندسازی فرآیند استخراج اطلاعات بپردازند. با این حال، ویژگی‌های داده‌ها می‌توانند به طور قابل توجهی بر پیچیدگی مشکلات در پیش‌بینی ناشی از هم‌پوشانی‌ها در مقادیر ویژگی‌ها، تفکیک‌پذیری داده‌ها و نفرین ابعاد^۱ تأثیر بگذارند (چارتی و همکاران^{۴۰}، ۲۰۲۲). روش تحقیق پیشنهادی در این مطالعه با بررسی کامل اثرات تکنیک‌های مهندسی ویژگی، از جمله تبدیل ویژگی و انتخاب ویژگی، بر عملکرد پیش‌بینی

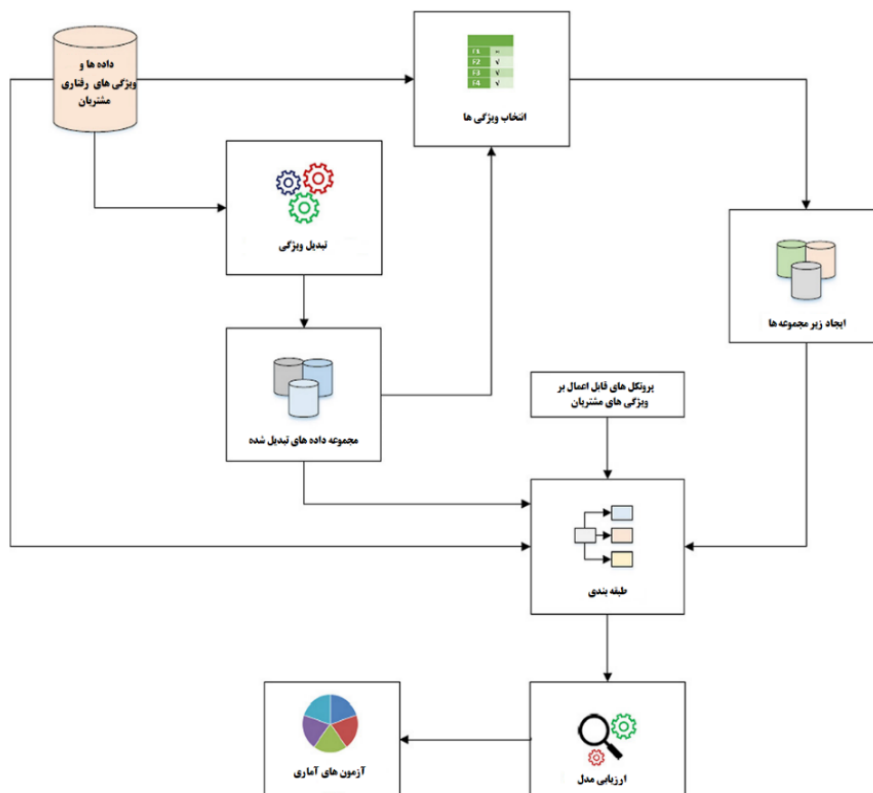
۱. نفرین ابعاد به پدیده‌های گوناگونی گفته می‌شود که هنگام تحلیل و ساماندهی داده‌ها در فضاهای بسیار بالا (اغلب با صدها یا هزاران بعد) روی می‌دهند، ولی نه در محیط‌های با ویژگی‌های بسیار پایین، مانند فضای فیزیکی. سه‌بعدی، که در زندگی روزمره احساس می‌گردد.

روش‌های مختلف طبقه‌بندی یادگیری ماشین، بر این مشکل غلبه می‌کند. برای دستیابی به دقت بالا، روش‌های طبقه‌بندی به نمایش‌های مناسب برای داده‌های ورودی خود نیاز دارند و تبدیل داده‌ها چنین داده‌هایی را برای کار داده شده مناسب‌تر می‌سازد. در داخل نیز تحقیقات متعددی در این حوزه صورت گرفته است.

سیفی و همکاران (۱۴۰۲)؛ به مدل‌سازی و شبیه‌سازی مدیریت الکترونیکی ارتباط با مشتریان با هدف مطالعه میزان وفاداری آن‌ها (مطالعه موردی: بانک تجارت) پرداختند. نتایج پژوهش نشان می‌دهد که بانک تجارت برای بقا در شرایط رقابتی بازار فعلی، بایستی به سمتی حرکت کند که مطلوبیت هر چهار عامل، یعنی بانکداری الکترونیک، پشتیبانی مشتری، امنیت و خدمات مازاد، افزایش پیدا کند. شناخت و درک صحیح از متغیرهای سیستم و نحوه اثرگذاری آن‌ها بر کل سیستم بسیار مهم است. استفاده از روش سیستم دینامیک این امکان را به‌خوبی فراهم می‌کند. قبولی و همکاران (۱۴۰۲)، به بررسی کشف دانش حاکم بر ویژگی‌های جمعیت‌شناختی مشتریان در انتخاب بانک‌ها با استفاده از قوانین انجمنی در داده کاوی پرداختند. این پژوهش بر اساس استاندارد CRISP-DM انجام گرفته و داده‌ها با پرسشنامه جمع‌آوری شده و نتایج آن به‌صورت بانک اطلاعاتی به یک نود منبع تبدیل و سپس با استفاده از نرم‌افزار spss modeler قوانین انجمنی مربوط به رفتار مشتریان آن بانک استخراج شده است. قوانین استخراج نشان‌دهنده تغییرها، چه تأثیری در سایر عوامل و در نهایت در دستیابی به اهداف دارند. خاتمی و فیروزآبادی (۱۳۹۷)؛ در تحقیق خود پنج ویژگی هریک از مشتریان شامل آخرین زمان مراجعه، تعداد تراکنش، مبلغ سپرده‌گذاری، مبلغ وام و مانده معوقات وام‌ها در طول یک سال فعالیت از پایگاه داده بانک استخراج شد و به کمک الگوریتم کامیانتگین، مشتریان خوشه‌بندی شدند. سپس یک مدل چندهدفه تخصیص خدمات بانک به هرکدام از خوشه‌ها طراحی گردید. اهداف مدل طراحی شده افزایش میزان رضایت مشتریان، کاهش هزینه‌ها و کاهش ریسک تخصیص خدمات است. با توجه به آنکه مسئله دارای یک راه حل بهینه نبوده و هر یک از ویژگی‌های مشتری دارای یک تابع توزیع احتمالی‌اند، برای حل از شبیه‌سازی استفاده شد. جهت تعیین جواب نزدیک به بهینه از الگوریتم تبرید در ساخت جواب‌های همسایه استفاده شد و مدل شبیه‌سازی اجرا گردید. نتایج به‌دست آمده بهبود قابل‌توجهی نسبت به وضعیت فعلی را نشان داد.

روش پژوهش

این تحقیق از نوع تحقیقات کاربردی و اکتشافی است. در این مطالعه در حدود ۹۸۶۰۴ مشتری بوده و با استفاده از نرم افزارهای ووکا و متلب اقدام به برآورد مدل های تحقیق شده است. در این تحقیق از ۴ مدل داده کاوی، فازی نوع ۱، ۲ و ۳؛ مدل های میانگین گیری بیزین غیر خطی بهره گرفته شده است. فرآیند انجام تحقیق در (نمودار ۱)، ترسیم شده است.



نمودار ۱. فرآیند انجام پژوهش

بر اساس نتایج (نمودار ۱)، ابتدا داده های مرتبط با مشتریان استخراج شده است. سپس بر اساس ویژگی های مشتریان اقدام به طبقه بندی مشتریان نموده و بر اساس آزمون های آماری میزان صحت طبقه بندی مشتریان مورد آزمون قرار گیرد (جدول ۱).

جدول ۱. مدل‌های کاربردی در تحقیق

ردیف	مدل	تعریف	کاربرد
طبقه‌بندی مشتریان	K-Means	کی-میانگین	یک الگوریتم تکراری است که مجموعه بدون برچسب را به k خوشه‌های مختلف تقسیم می‌کند به گونه‌ای که هر مجموعه‌ای فقط متعلق به یک گروه است که دارای ویژگی‌های مشابه است.
طبقه‌بندی مشتریان	- Means C	سی-میانگین	در این رویکرد از خاصیت فازی جهت طبقه‌بندی داده‌ها بهره گرفته می‌شود.
لحاظ ناطمینانی در اطلاعات استخراجی	Type-3 fuzzy	فازی نوع ۳	لحاظ نمودن مبحث ناطمینانی در امتیازدهی عوامل موثر بر نوسانات قیمت رمز ارزها توسط خبرگان
شناسایی بهترین ویژگی‌های مشتریان بانک (رویکردهای مدل‌ساز)	BMA	مدل‌های میانگین‌گیری بیزین	انتخاب مهم‌ترین ویژگی‌های مشتریان بانک
	TVP-DMA	مدل‌های میانگین‌گیری پویا پارامتر متغیر زمان	انتخاب مهم‌ترین ویژگی‌های مشتریان بانک
	TVP-DMS	مدل‌های میانگین‌گیری انتخابی پارامتر متغیر زمان	انتخاب مهم‌ترین ویژگی‌های مشتریان بانک

یافته‌ها

مدل‌های پارامتر متغیر در طول زمان (TVP)، روش‌های فضا حالت (مانند فیلتر کالمن) را به کار می‌گیرند که این موضوع، عموماً در تحقیقات تجربی اقتصاد کلان در راستای تجزیه و تحلیل ساختاری و پیش‌بینی استفاده می‌شود. چنانچه مجموعه بزرگی از داده‌ها به منظور پیش‌بینی متغیرهای کلان اقتصادی استفاده گردد، مدل‌های TVP تمایل به بیش‌برازشی در داخل نمونه دارند؛ لذا عملکرد پیش‌بینی ضعیفی در خارج از نمونه خواهند داشت. برای تصحیح این کاستی‌ها در مدل‌های TVP از مدل‌های DMS و DMA، استفاده شده است (گوپتا و همکاران، ۲۰۱۴).

با توجه به اینکه رویکرد $TVP - DMA$ تابع (α, λ) می باشد. در جدول زیر نتایج بر اساس (α, λ) مختلف ارائه شده است.

جدول ۲. معیارهای عملکرد پیش‌بینی در افق‌های پیش‌بینی مختلف

	پیش‌بینی ۵ درصد از داده‌ها					
	LOG(PL)	MAFE	MSFE	MAPE	FEV	Bias
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.99)$	68.958	0.071	0.009	0.187	0.009	0.017
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.95)$	76.309	0.062	0.007	0.183	0.007	0.014
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.90)$	78.001	0.057	0.006	0.168	0.006	0.013
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.99)$	69.739	0.076	0.011	0.191	0.010	0.018
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.95)$	80.483	0.067	0.008	0.169	0.008	0.011
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.90)$	100.298	0.053	0.006	0.152	0.006	0.015
$TVP - AR(1) - X DMA(\alpha = 0.99, \lambda = 1)$	66.599	0.073	0.010	0.194	0.009	0.016
$TVP - AR(1) - X DMA(\alpha = 0.95, \lambda = 1)$	71.045	0.067	0.008	0.221	0.007	0.023
$TVP - AR(1) - X BMA(\alpha = \lambda = 1)$	109.698	0.014	0.002	0.106	0.021	0.005
پیش‌بینی ۱۰ درصد از داده‌ها						
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.99)$	65.321	0.074	0.010	0.183	0.010	0.020
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.95)$	72.154	0.062	0.007	0.171	0.007	0.015
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.90)$	73.367	0.057	0.006	0.160	0.006	0.014
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.99)$	65.415	0.079	0.011	0.187	0.011	0.020
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.95)$	75.078	0.068	0.008	0.167	0.008	0.009
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.90)$	92.045	0.057	0.007	0.161	0.007	0.009
$TVP - AR(1) - X DMA(\alpha = 0.99, \lambda = 1)$	63.036	0.074	0.010	0.185	0.009	0.015
$TVP - AR(1) - X DMA(\alpha = 0.95, \lambda = 1)$	68.714	0.066	0.007	0.194	0.007	0.021
$TVP - AR(1) - X BMA(\alpha = \lambda = 1)$	93.295	0.016	0.003	0.099	0.002	0.014
پیش‌بینی ۲۰ درصد از داده‌ها						
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.99)$	61.514	0.076	0.010	0.516	0.010	0.010
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.95)$	68.141	0.062	0.007	0.378	0.007	0.012
$TVP - AR(1) - X DMA(\alpha = \lambda = 0.90)$	69.137	0.056	0.006	0.298	0.006	0.013
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.99)$	59.502	0.080	0.011	0.519	0.011	0.009
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.95)$	71.694	0.071	0.010	0.432	0.009	0.008
$TVP - AR(1) - X DMS(\alpha = \lambda = 0.90)$	85.258	0.061	0.008	0.402	0.008	0.011
$TVP - AR(1) - X DMA(\alpha = 0.99, \lambda = 1)$	63.177	0.073	0.009	0.534	0.009	0.010
$TVP - AR(1) - X DMA(\alpha = 0.95, \lambda = 1)$	68.197	0.062	0.007	0.445	0.006	0.012
$TVP - AR(1) - X BMA(\alpha = \lambda = 1)$	78.255	0.016	0.002	0.074	0.005	0.002

بر اساس نتایج مدل BMA در تمامی حالت‌ها از عملکرد مطلوب‌تری برخوردار است. در نتیجه در ادامه به بررسی نتایج مدل BMA پرداخته می شود. در این رویکرد ابتدا تمامی

حالت‌های ممکن حضور متغیرهای توضیحی بر روی متغیر وابسته رگرسیون می‌گردد. در این روش چندین نکته حائز اهمیت است. اول اینکه یک متغیر در تمامی مدل‌های ممکن حضور ندارد. نکته دوم اینکه لزوماً متغیر مذکور در تمامی مدل‌هایی که حضور دارد تأثیر معناداری بر متغیر وابسته ندارد؛ بر این اساس نسبت تعداد مدل‌هایی که متغیر مذکور معنادار شده به تعداد مدل‌هایی که حضور دارد؛ شاخصی جهت حضور متغیر مذکور در مدل بهینه است؛ سوم اینکه با افزایش تعداد متغیرها محاسبه تمامی حالت‌ها امکان‌پذیر نیست. در نتیجه بر اساس دیدگاه سالای و مارتین از یک تعداد برآورد به بعد، نسبت حضور معنادار یک متغیر به تمامی حالت‌ها، به سمت یک عدد مشخص میل می‌نماید و در نتیجه نیازی به برآورد تمامی حالت‌ها نمی‌باشد.

در نهایت نیاز به یک آستانه تصمیم‌گیری جهت حذف متغیرها وجود دارد؛ برای تعیین حد بهینه از نسبت k تقسیم بر کل متغیرها بهره گرفته خواهد شد (k تعداد متغیرهای پیشنهادی است که از دیدگاه محقق بالاترین تأثیر را بر متغیر وابسته دارند). این k تجربی بوده و بر اساس دیدگاه محقق انتخاب می‌شود.

با توجه به تعداد متغیرها ۲۲۲ مدل وجود دارد که بیش از ۴،۱۹۴،۳۰۴ مدل رگرسیونی برآورد می‌گردد. بر این اساس در این تحقیق باید بخشی از این تعداد مدل برآورد گردد؛ به گونه‌ای که نتایج آن قابلیت تعمیم به کل مدل‌های برآوردی را داشته باشد (شرط همگرایی). (جدول ۳). جهت شناسایی مهمترین متغیرهای موثر بر تخصیص خدمات بانکی با استفاده از رویکرد میانگین‌گیری بیزین دو شرط اصلی باید اعمال گردد.

(۱) افزایش احتمال پسین هر متغیر نسبت به احتمال پیشین

(۲) بالا بودن سطح احتمال پسین از سطح آستانه تعریف شده («سطح آستانه اولیه=۵ تقسیم بر ۲۲=۰/۲۲۷۲»). نتایج این بخش در جدول زیر ارائه شده است.

جدول ۳. مرحله اول فرآیند نمونه گیری و محاسبات با فرض $\bar{K} = 5$

نمونه اول شامل ۲ میلیون رگرسیون		نمونه اول شامل ۱ میلیون رگرسیون		متغیر
احتمال پسین	ضریب پسین	احتمال پیشین	ضریب پیشین	
0.321	0.132	0.207	0.002	ویژگی‌های شخصی - جنسیت
0.292	0.003	0.173	0.004	ویژگی‌های شخصی - سن
0.078	0.176	0.035	0.319	تاریخ افتتاح حساب
0.181	0.034	0.170	0.029	نوع حساب
0.188	0.051	0.107	0.080	تاریخ آخرین مراجعه مشتری
0.155	0.209	0.118	0.428	وضعیت حساب
0.134	0.788	0.107	0.147	مدت زمان استفاده از حساب
0.162	0.068	0.109	0.111	امتیاز حساب
0.569	0.127	0.445	0.417	موجودی حساب
0.146	0.023	0.070	0.024	بازه زمانی بین نخستین و آخرین بازدید مشتری
0.247	0.400	0.122	0.000	تعداد تراکنش‌ها در طول بازه زمانی
0.141	0.007	0.122	0.006	وضعیت دسته چک
0.261	0.127	0.202	0.599	تعداد تسهیلات
0.162	0.022	0.130	0.039	نوع تسهیلات
0.269	0.718	0.179	0.692	تعداد معوقات
0.728	0.000	0.509	0.000	حجم تراکنش‌های مالی
0.185	0.000	0.152	0.000	تعداد سپرده‌ها
0.830	0.188	0.638	0.059	مجموع مانده سپرده‌ها
0.412	0.955	0.162	0.489	میزان تسهیلات جاری
0.395	0.044	0.190	0.189	مجموع مانده کل تسهیلات جاری
0.240	0.025	0.173	0.015	مجموع کل معوقات
0.149	0.546	0.123	0.129	تاریخ آخرین مطالبات

منبع: محاسبات محقق

در مرحله‌ی اول با استفاده از شروط دو گانه فوق، جهت تعیین نمودن متغیرهای غیر شکننده، ۱۱ متغیر انتخاب شدند. یعنی ۱۱ متغیر مقدار احتمال پسین بیش‌تری نسبت به احتمال پیشین داشتند و این ۱۱ متغیر سطح احتمال پسین بالاتر از سطح

آستانه ۰/۲۲۷۲ داشتند. در ادامه تمامی مراحل انجام شده در مرحله اول را در مرحله دوم بر روی ۱۱ متغیر باقیمانده اعمال خواهیم نمود. در مرحله دوم نیز ابتدا یک نمونه شامل ۱ هزار رگرسیون بر روی ۱۱ متغیر منتخب اعمال شده و محاسبات ضرایب و احتمالات پسین صورت گرفته است. در ادامه با اعمال دو شرط مذکور «سطح آستانه ثانویه = ۵ تقسیم ۱۱ = ۰/۴۵۴۵»؛ مهم ترین متغیرهای موثر بر خوشه بندی مشتریان شناسایی خواهند شد. نتایج در (جدول ۴)، می توان دید.

جدول ۴. مرحله دوم فرآیند نمونه گیری و محاسبات با فرض $K = 5$

متغیر		نمونه اول شامل ۱ هزار رگرسیون		نمونه اول شامل ۲ هزار رگرسیون	
		ضریب پیشین	احتمال پیشین	ضریب پسین	احتمال پسین
ویژگی های شخصیتی -جنسیت		0.002	0.246	0.157	0.382
ویژگی های شخصیتی سن		0.005	0.206	0.004	0.347
موجودی حساب		0.496	0.530	0.151	0.677
تعداد تراکنش ها در طول بازه زمانی		0.000	0.145	0.476	0.294
تعداد تسهیلات		0.713	0.240	0.151	0.311
تعداد معوقات		0.823	0.213	0.854	0.320
حجم تراکنش های مالی		0.145	0.606	0.224	0.866
مجموع مانده سپرده ها		0.070	0.759	0.224	0.988
میزان تسهیلات جاری		0.582	0.193	1.136	0.441
مجموع مانده کل تسهیلات جاری		0.225	0.226	0.052	0.571
مجموع کل معوقات		0.018	0.206	0.030	0.286

منبع: محاسبات محقق

تعیین خوشه بهینه

پس از تعیین شاخص های موثر بر خوشه بندی مشتریان لازم است اقدام به تعیین تعداد خوشه ها گردد. تعیین تعداد خوشه بهینه، میانگین ضریب سیلوئیت با استفاده از دو الگوریتم K-Means و C-Means فازی، در (جدول ۵)، ارائه شده است.

جدول ۵. میانگین ضریب سیلوئت با استفاده از الگوریتم K-Means و C - Means فاز ۱

شماره وقفه	k-means	c - means فاز ۱
خوشه ۲	۰/۸۷۰۳	۰/۸۷۲۳
خوشه ۳	۰/۷۷۴۵	۰/۷۹۳۵
خوشه ۴	۰/۷۳۵۵	۰/۷۴۸۲
خوشه ۵	۰/۷۱۰۶	۰/۷۰۰۳
خوشه ۶	۰/۶۹۰۹	۰/۶۸۳۹
خوشه ۷	۰/۶۷۰۵	۰/۶۳۹۴
خوشه ۸	۰/۵۹۹۵	۰/۶۱۸۳
خوشه ۹	۰/۶۳۴۲	۰/۵۹۳۴
خوشه ۱۰	۰/۶۹۸۲	۰/۵۵۸۲
خوشه ۱۱	۰/۷۱۵۰	۰/۵۲۵۵
خوشه ۱۲	۰/۷۸۳۵	۰/۵۱۸۴
خوشه ۱۳	۰/۸۲۴۵	۰/۷۳۹۴
خوشه ۱۴	۰/۷۲۲۲	۰/۷۵۷۳
خوشه ۱۵	۰/۶۹۳۲	۰/۸۲۸۴
خوشه ۱۶	۰/۶۵۳۳	۰/۹۲۰۳
خوشه ۱۷	۰/۶۲۸۶	۰/۹۰۴۳
خوشه ۱۸	۰/۶۰۷۴	۰/۸۴۳۲
خوشه ۱۹	۰/۵۸۷۳	۰/۷۴۴۷
خوشه ۲۰	۰/۵۵۳۴	۰/۶۴۶۲

منبع: محاسبات محقق

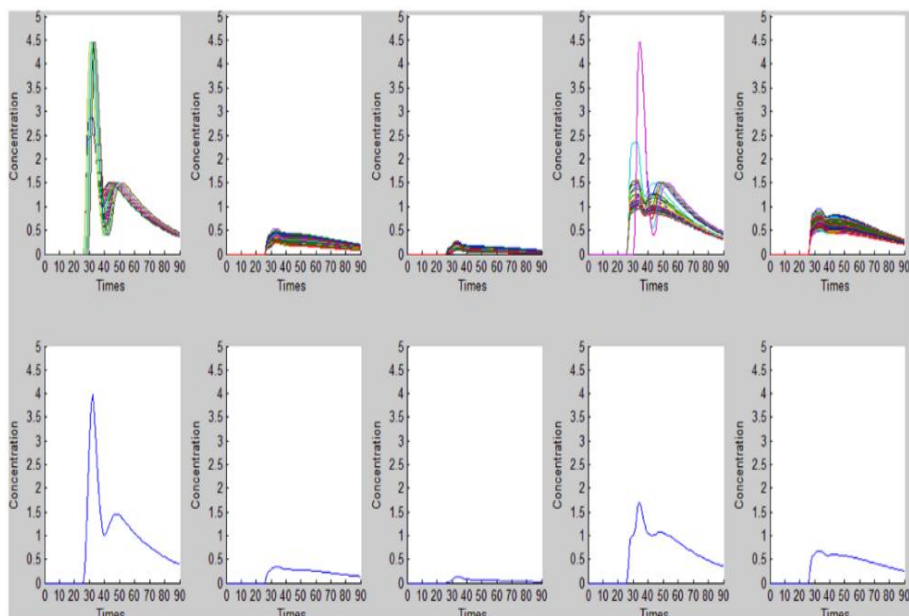
بر اساس نتایج در حالت K-Means ۱۳ خوشه و در حالت C - Means فاز ۱ ۱۶ خوشه به عنوان خوشه بهینه انتخاب شده است. در (جدول ۶)، میزان خطای هر یک از رویکردهای فوق ارائه شده است.

جدول ۶. مقایسه دقت مدل K-Means و C - Means فاز ۱

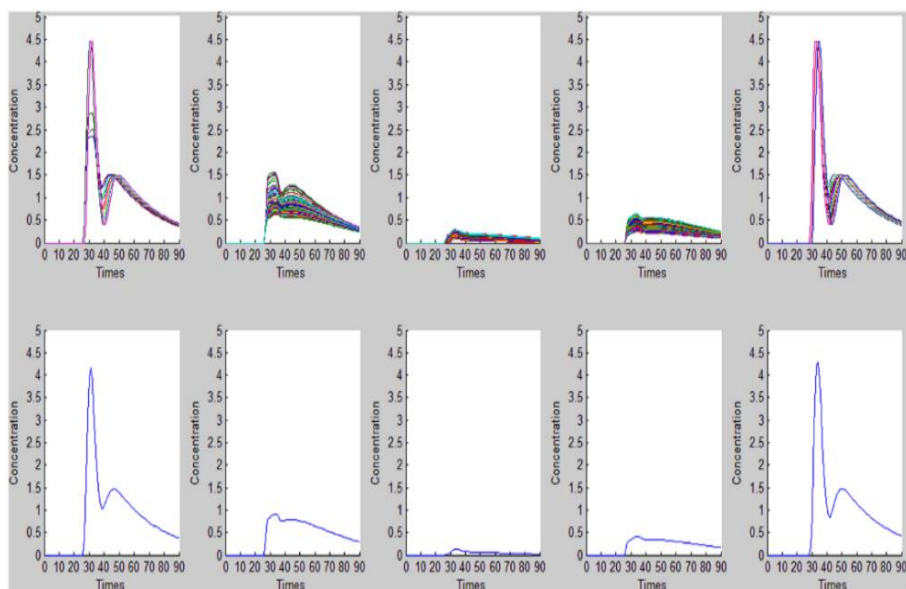
شماره وقفه	k-means	c - means فاز ۱
RMSE	۰/۱۳۸۴	۰/۰۹۳۶

منبع: محاسبات محقق

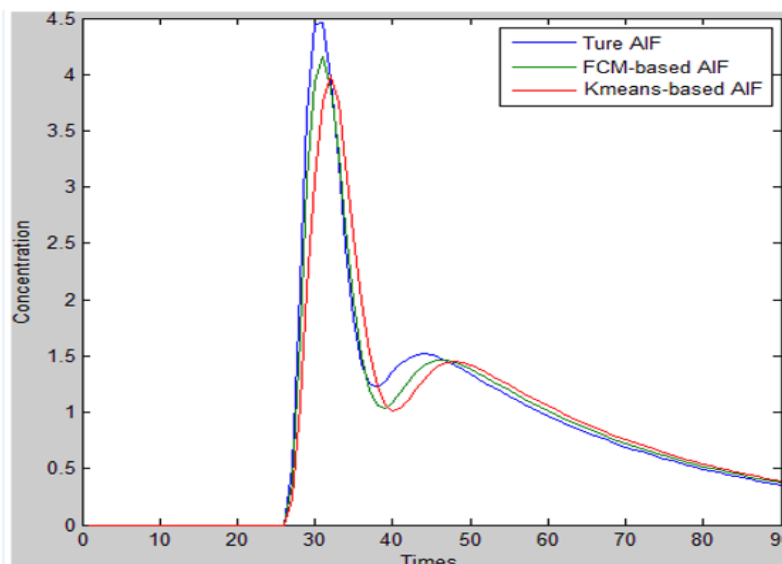
بر اساس نتایج خطای مدل دقت C - Means فاز ۱ از دقت مدل K-Means بالاتر است (نمودار ۲) و (نمودار ۳).



نمودار ۲. نتایج C -Means فازی در خوشه‌های مختلف منبع: محاسبات محقق



نمودار ۳. نتایج K -Means در خوشه‌های مختلف منبع: محاسبات محقق



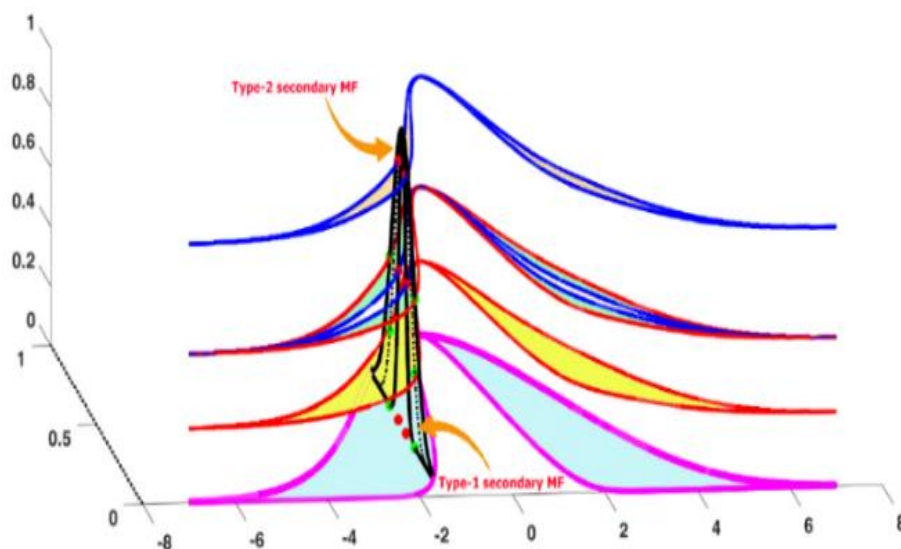
نمودار ۴. مقایسه نتایج مدل K -Means و C -Means فازی

منبع: محاسبات محقق

بر اساس (نمودار ۴)، مشاهده می‌گردد که مدل C -Means فازی توانایی بالاتری در خوشه‌بندی داده‌ها را بر اساس داده واقعی دارد. در نتیجه در ادامه مابقی برآوردهای تحقیق بر اساس این رویکرد ادامه خواهد یافت. قبل از برآورد اثرگذاری متغیرهای موثر بر سودآوری بانک لازم است از کارایی انواع مدل های فازی در C -Means اطمینان حاصل کنیم.

سه رویکرد در حوزه فازی کردن وجود دارد که به فازی نوع اول؛ دوم و سوم شهرت دارند. رویکرد فازی نوع سه در اکثر تحقیقات مورد تأیید قرار گرفته است. در این تحقیق نیز از هر سه نوع سیستم فازی بهره گرفته خواهد شد. سیستم‌های منطق فازی نوع ۳ می‌توانند سطح بالاتری از عدم قطعیت‌ها را در مقایسه با نمونه‌های نوع ۲ ارائه کنند؛ زیرا عضویت ثانویه و عدم قطعیت‌های بالا و پایین در مجموعه فازی نوع ۳ مقادیر واضح نیستند؛ اما مجموعه‌های فازی هستند. برای مشاهده بهتر تفاوت این دو رویکرد، یک برش افقی در مجموعه‌های فازی نوع ۳ و نوع ۲ را در نظر بگیرید؛ همانطور که در (نمودار ۵)، نشان داده شده است، مشاهده می‌شود که مرزهای عدم قطعیت (بالا و

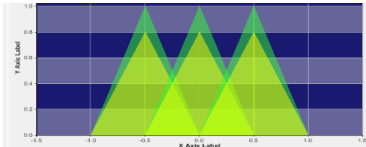
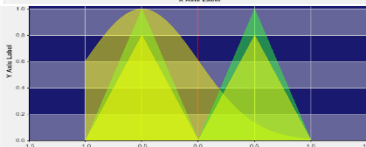
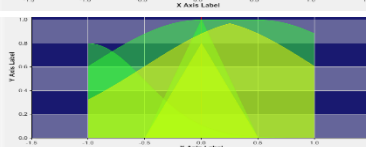
پایین عضویت‌های اولیه) در مجموعه فازی نوع ۳ مقادیر واضحی نیستند؛ اما مجموعه‌های فازی را نیز بازه‌ای می‌کنند.



نمودار ۵. برش افقی برای منطق فازی نوع ۳

در این مقاله، فرم غیر تکی سیستم‌های منطق فازی پایه نوع ۳ در یک طرح تشخیص عیب فرموله شده است. در (جدول ۷)، مقایسه‌ای از فازی‌های نوع ۱ و ۲ و ۳ ارائه شده است. با توجه به اینکه در فازی نوع دو نسبت به فازی نوع یک و در فازی نوع سه نسبت به فازی نوع یک و دو عدم اطمینان بیشتری پوشش داده می‌شود، استفاده از توابع ورودی که بتواند دامنه نااطمینانی بالاتری را پوشش دهد. بر این اساس توابع ورودی فازی از نوع ترکیبی مثلثی-گوسی انتخاب گردید تا بالاترین پوشش دامنه نااطمینانی را دارا است.

جدول ۷. مقایسه نتایج رویکرد فازی نوع ۱، ۲ و ۳

کارایی میانگین			نوع تابع	فرم تابع
فازی نوع ۳	فازی نوع ۲	فازی نوع ۱		
۰/۷۱۲	۰/۶۵۸	۰/۶۱۱	توابع ورودی فازی از نوع مثلثی	
۰/۷۹۴	۰/۷۲۱	۰/۶۴۳	توابع ورودی فازی از نوع ترکیبی مثلثی-نیمه گاوسی	
۰/۹۴۸	۰/۸۱۳	۰/۶۹۴	توابع ورودی فازی از نوع ترکیبی مثلثی-گاوسی	
بین فازی نوع ۱ و نوع ۲ و نوع ۳	بین فازی نوع ۱ و نوع ۳	بین فازی نوع ۲ و نوع ۳	بین فازی نوع ۱ و نوع ۲	
۲۲/۱۹	۱۳/۲۴	۸/۱۱	۹/۷۸	
تفاوت معنادار میان میانگین‌ها وجود دارد.	تفاوت معنادار میان میانگین‌ها وجود دارد.	تفاوت معنادار میان میانگین‌ها وجود دارد.	تفاوت معنادار میان میانگین‌ها وجود دارد.	

آماره T زوجی (دو میانگین) و F تجزیه واریانس (بیش از دو میانگین)

منبع: محاسبات محقق

بر اساس نتایج تفاوت معناداری میان کارایی میانگین سه رویکرد فازی وجود دارد؛ در نتیجه فازی نوع ۳ جهت اعمال شده است. در ادامه نتایج بر اساس فازی نوع ۳ ارائه شده است.

طبقه‌بندی مشتریان

مرحله یک: تعیین وزن شاخص مدل LRFM با استفاده از تکنیک فرآیند تحلیل سلسله

مراتبی گروهی: در این تحقیق برای وزن دهی شاخص های مدل LRFM از روش تحلیل سلسله مراتبی گروهی و با مقایسه زوجی شاخص ها، ضریب اهمیت هر یک از آن ها محاسبه گردید. برای این منظور از نظرات ۵ مدیر ارشد بانک استفاده شده است (جدول ۸).

جدول ۸. تعیین وزن شاخص های مدل

بردار ویژه	F	M	R	L	
(L)	2.208	3.354	2.787	1	(L)
(R)	0.336	0.530	1	0.359	(R)
(M)	1.765	1	1.888	0.283	(M)
(F)	1	0.566	2.980	1.888	(F)

منبع: محاسبات محقق

نرخ ناسازگاری مقایسه های انجام شده نیز برابر ۰/۰۸۶ بدست آمده است که چون کوچک تر از ۰/۱ می باشد؛ بنابراین می توان به مقایسه های انجام شده اطمینان کرد. مرحله دو: تعیین میانگین کل مشتریان: در این مرحله میانگین هر یک از شاخص های L,R,F,M برای داده های استخراج شده در فاز یک بدست می آوریم که میانگین ۴ شاخص برای ۹۸۶۰۴ مشتری از خروجی بیزین باید محاسبه گردد. مرحله سه: گروه بندی مشتریان: در این مرحله مشتریان در قالب ۱۶ گروه طبقه بندی می گردند. از آنجا که هر شاخص می تواند دو حالت به خود گیرد (بالتر از میانگین و یا پایین تر از میانگین) بر این اساس و با توجه به تعداد ویژگی ها (۴ شاخص موجودی حساب؛ مجموع مانده سپرده ها؛ مجموع مانده کل تسهیلات جاری و حجم تراکنش های مالی) دارای مزایا برای بانک است؛ تعداد کل حالات ممکن ۱۶ حالت می باشد نوع گروه بر این اساس ساخته می شود که برای تک تک مشتریان مقدار هر یک از شاخص های با میانگین کل که در مرحله قبل بدست آمده مقایسه می گردد که آیا پایین تر از میانگین است یا بالاتر از میانگین. در حالتی که مقدار مشاهده شده از میانگین در کل داده ها بالاتر باشد و در صورتی که مقدار مشاهده شده از میانگین در کل داده ها پایین تر از میانگین در کل داده ها باشد. این مرحله با کمک نرم افزار وکا انجام می گردد.

مرحله چهار: محاسبه میانگین مشتریان هر گروه: در این مرحله مقادیر نرمال شده شاخص های برای مشتریان قرار گرفته در هر گروه از فاز ۲ استخراج شده و میانگین مقادیر نرمال ویژگی ها در هر گروه از مشتریان تعیین می گردد. داده های استخراج شده از این مرحله در (جدول ۹)، نشان داده شده است. میانگین ویژگی ها در (جدول ۹)، ارائه شده اند.

جدول ۹. میانگین مقادیر ۴ شاخص برای هر گروه و نوع گروه های مشتریان

سناریو	رویکرد			
1	M پایین تر از میانگین	F پایین تر از میانگین	R پایین تر از میانگین	L پایین تر از میانگین
2	M بالاتر از میانگین	F پایین تر از میانگین	R پایین تر از میانگین	L پایین تر از میانگین
3	M پایین تر از میانگین	F بالاتر از میانگین	R پایین تر از میانگین	L پایین تر از میانگین
4	M پایین تر از میانگین	F بالاتر از میانگین	R پایین تر از میانگین	L پایین تر از میانگین
5	M پایین تر از میانگین	F پایین تر از میانگین	R پایین تر از میانگین	L بالاتر از میانگین
6	M بالاتر از میانگین	F بالاتر از میانگین	R پایین تر از میانگین	L پایین تر از میانگین
7	M بالاتر از میانگین	F پایین تر از میانگین	R بالاتر از میانگین	L پایین تر از میانگین
8	M بالاتر از میانگین	F پایین تر از میانگین	R پایین تر از میانگین	L بالاتر از میانگین
9	M پایین تر از میانگین	F بالاتر از میانگین	R بالاتر از میانگین	L پایین تر از میانگین
10	M پایین تر از میانگین	F بالاتر از میانگین	R پایین تر از میانگین	L بالاتر از میانگین
11	M پایین تر از میانگین	F پایین تر از میانگین	R بالاتر از میانگین	L پایین تر از میانگین
12	M بالاتر از میانگین	F بالاتر از میانگین	R بالاتر از میانگین	L پایین تر از میانگین
13	M بالاتر از میانگین	F بالاتر از میانگین	R پایین تر از میانگین	L بالاتر از میانگین
14	M بالاتر از میانگین	F پایین تر از میانگین	R بالاتر از میانگین	L بالاتر از میانگین
15	M پایین تر از میانگین	F بالاتر از میانگین	R بالاتر از میانگین	L بالاتر از میانگین
16	M بالاتر از میانگین	F بالاتر از میانگین	R بالاتر از میانگین	L بالاتر از میانگین

منبع: محاسبات محقق

تذکر: رنگ سبز: بیانگر مشتریان وفادار؛ رنگ نارنجی: بیانگر مشتریان نسبتاً وفادار؛ رنگ زرد: بیانگر مشتریان غیر وفادار؛ رنگ قرمز: بیانگر مشتریان غیر واقعی

مرحله پنجم: محاسبه و رتبه بندی و اولویت بندی CLV گروه های مشتریان: در این قسمت CLV مشتریان قرار گرفته در هر گروه محاسبه و رتبه بندی و اولویت بندی می شوند. CLV براساس یک نرخ یکپارچه برای هر گروه محاسبه می شود. نرخ یکپارچه گروه زام عبارت است از:

$$CjI = wLCjL + wRCjR + wFCjF + wMCjM$$

که در آن CjM, CjF, CjR, CjL میانگین مقادیر نرمال شده LRFM برای مشتریان قرار گرفته در هر گروه و wM, wF, wR, wL وزن ویژگی‌ها است که به روش فرآیند AHP گروهی محاسبه می‌شود.

مرحله ششم: بخش بندی: در این مرحله ۱۶ گروه شناسایی شدند، شناسایی هر گروه بستگی به میزان شباهت رفتاری هر گروه با یکدیگر است.

خوشه یک مشتریان وفادار: مشتریان خوشه یک مشتریانی هستند که دارای شاخص‌های زیر باشند.

موجودی حساب (+)؛

مجموع مانده سپرده‌ها (+)؛

مجموع مانده کل تسهیلات جاری (+)؛

حجم تراکنش‌های مالی (+) باشند.

در نتیجه سناریوهایی که بالاتر از سه یا چهار علامت موافق را داشته باشند مشتریان وفادار و بالقوه وفادار خواهند بود (جدول ۱۰).

جدول ۱۰. بخش‌بندی گروه مشتریان وفادار و بالقوه وفادار

درصد	سناریو
۸/۷ درصد	شماره ۱۲ M بالاتر از میانگین؛ F بالاتر از میانگین؛ R بالاتر از میانگین؛ L پایین‌تر از میانگین
۴/۷ درصد	شماره ۱۳ (M بالاتر از میانگین؛ F بالاتر از میانگین؛ R پایین‌تر از میانگین؛ L بالاتر از میانگین)
۳/۴ درصد	شماره ۱۴ (M بالاتر از میانگین؛ F پایین‌تر از میانگین؛ R بالاتر از میانگین؛ L بالاتر از میانگین)
۲/۷ درصد	شماره ۱۵ (M پایین‌تر از میانگین؛ F بالاتر از میانگین؛ R بالاتر از میانگین؛ L بالاتر از میانگین)
۱/۶ درصد	شماره ۱۶ M بالاتر از میانگین؛ F بالاتر از میانگین؛ R بالاتر از میانگین؛ L بالاتر از میانگین
۲۱/۱ درصد	جمع

منبع: محاسبات محقق

خوشه دو مشتریان نسبتاً وفادار: مشتریان خوشه دو مشتریانی هستند که دارای حداقل دو علامت موافق باشند (جدول ۱۱).

جدول ۱۱. بخش‌بندی گروه مشتریان نسبتاً وفادار

سناریو	درصد
شماره ۶ (M بالاتر از میانگین؛ F بالاتر از میانگین؛ R پایین‌تر از میانگین؛ L پایین‌تر از میانگین)	۵/۶ درصد
شماره ۷ (M بالاتر از میانگین؛ F پایین‌تر از میانگین؛ R بالاتر از میانگین؛ L پایین‌تر از میانگین)	۴/۱ درصد
شماره ۸ (M بالاتر از میانگین؛ F پایین‌تر از میانگین؛ R پایین‌تر از میانگین؛ L بالاتر از میانگین)	۱۰/۲ درصد
شماره ۹ (M پایین‌تر از میانگین؛ F بالاتر از میانگین؛ R بالاتر از میانگین؛ L پایین‌تر از میانگین)	۵/۹ درصد
شماره ۱۰ (M پایین‌تر از میانگین؛ F بالاتر از میانگین؛ R پایین‌تر از میانگین؛ L بالاتر از میانگین)	۸/۴ درصد
جمع	۳۴/۲ درصد

منبع: محاسبات محقق

خوشه سوم مشتریان غیر وفادار: مشتریان خوشه سوم مشتریانی هستند که یک علامت موافق (حالت مطلوب) در سناریوهایی آن‌ها وجود داشته باشد (جدول ۱۲).

جدول ۱۲. بخش‌بندی گروه مشتریان غیر وفادار

سناریو	درصد
شماره ۲ (M بالاتر از میانگین؛ F پایین‌تر از میانگین؛ R پایین‌تر از میانگین؛ L پایین‌تر از میانگین)	۴/۶ درصد
شماره ۳ (M پایین‌تر از میانگین؛ F بالاتر از میانگین؛ R پایین‌تر از میانگین؛ L پایین‌تر از میانگین)	۹/۳ درصد
شماره ۴ (M پایین‌تر از میانگین؛ F بالاتر از میانگین؛ R پایین‌تر از میانگین؛ L پایین‌تر از میانگین)	۸/۳ درصد
شماره ۵ (M پایین‌تر از میانگین؛ F پایین‌تر از میانگین؛ R پایین‌تر از میانگین؛ L بالاتر از میانگین)	۶/۶ درصد
شماره ۱۱ (M پایین‌تر از میانگین؛ F پایین‌تر از میانگین؛ R بالاتر از میانگین؛ L پایین‌تر از میانگین)	۷/۴ درصد
	۳۶/۲ درصد

منبع: محاسبات محقق

خوشه چهارم مشتریان غیر واقعی: مشتریان خوشه چهار مشتریانی هستند که دارای شاخص‌های زیر باشند. در نتیجه سناریوهایی که هیچ علامت موافق (حالت مطلوب) را نداشته باشند مشتریان غیر واقعی خواهند بود (جدول ۱۳).

جدول ۱۳. بخش‌بندی گروه مشتریان غیر واقعی

سناریو	درصد
شماره ۱ (M پایین‌تر از میانگین، F پایین‌تر از میانگین، R پایین‌تر از میانگین، L پایین‌تر از میانگین)	۸/۵ درصد
جمع	۸/۵ درصد

منبع: محاسبات محقق

فاز ۵: ارزیابی: مجموعه فعالیت‌هایی که در این بخش می‌بایست انجام شود عبارتند از: توابع تشخیصی برای پیش‌بینی و تشخیص مشتریان آینده استفاده می‌گردد. در این بخش ۱۶ گروه به وسیله چهار شاخص M, F, R, L شناسایی شدند؛ زیرا مقدار سطح معنی‌داری برای هر یک از ویژگی‌ها کمتر از ۰/۰۵ است (جدول ۱۴).

جدول ۱۴. آزمون برابری میانگین گروه‌ها برای هر یک از چهار شاخص مورد استفاده در تحلیل

Tests of Equality of Group Means					
Sig.	df2	df1	F آماره	Wilks' Lambda	
0.000	۱۰۴۹۹۰	15	33.096	0.847	L
0.000	۱۰۴۹۹۰	15	56.996	0.409	R
0.000	۱۰۴۹۹۰	15	56.445	0.395	F
0.000	۱۰۴۹۹۰	15	44.65	0.574	M

منبع: محاسبات محقق

مقادیر، مقادیر ویژه^{۴۱} بیانگر نسبت واریانسی است که به وسیله تحلیل توضیح داده شده است که در (جدول ۱۵)، نتایج حاصل و نیز مقادیر همبستگی کانونی^{۴۲} قابل مشاهده هستند.

جدول ۱۵. نتایج مقادیر ویژه و همبستگی کانونی برای مشتریان Eigenvalues

همبستگی کانونی	تجمعی %	انحراف %	مقدار ویژه	تابع
0.921	65.9	65.9	5.584	1
0.766	82.7	16.8	1.423	2
0.753	98.2	15.5	1.313	3
0.363	100.0	1.8	0.152	4

منبع: محاسبات محقق

اعداد حاصل از (جدول ۱۵)، بیانگر یک همبستگی نسبتاً بالا و در نتیجه تحلیل تشخیصی بین گروه‌های مختلف بر حسب چهار شاخص مقبول است.

برآورد توابع تشخیصی برای پیش‌بینی و تشخیص مشتریان با استفاده از تحلیل تشخیصی ۱۶ معادله شناسایی گردید. از این توابع برای پیش‌بینی و

تشخیص هر یک از مشتریان فعلی که جز ۹۸۶۰۴ مشتری نبوده‌اند، بهره گرفته می‌شود. تا زمانی که مقادیر شاخص‌های L, M, F, R برای هر مشتری با مقادیر میانگین کل مشتریان تفاوت زیادی نداشته باشد این توابع جهت پیش‌بینی و تشخیص وضعیت مشتریان استفاده می‌شود. مقادیر چهار شاخص مشتری جدید را در ۱۶ معادله برآوردی قرار داده و بالاترین تطابق در هر مرحله‌ای باشد، مشتری مورد نظر به همان گروه تعلق می‌گیرد. توابع تشخیصی گروه‌های مختلف در ذیل ارائه شده است:

گروه اول $Di = -27.054 - 0.104(Li) - 0.176(Ri) - 0.436(Mi) - 2.094E-8(Fi)$

گروه دوم $Di = 6.995 - 0.051(Li) - 0.046(Ri) - 0.089(Mi) + 2.374E-8(Mi)$

گروه سوم $Di = 9.143 - 0.020(Li) - 0.038(Ri) + 0.599(Mi) - 1.026E-9(Fi)$

گروه چهارم $Di = 12.874 - 0.042(Li) - 0.189(Ri) + 0.076(Mi) - 1.950E-9(Fi)$

گروه پنجم $Di = 18.483 + 0.115(Li) - 0.077(Ri) - 0.027(Mi) - 1.914E-9(Fi)$

گروه ششم $Di = 15.750 - 0.015(Li) - 0.038(Ri) + 0.849(Mi) + 1.292E-8(Fi)$

گروه هفتم $Di = 19.720 - 0.036(Li) + 0.247(Ri) - 0.137(Mi) + 2.269E-8(Fi)$

گروه هشتم $Di = -19.78 + 0.118(Li) - 0.077(Ri) - 0.028(Fi) - 2.022E-8(Fi)$

گروه نهم $Di = 21.483 - 0.063(Li) + 0.207(Ri) \pm 0.514(Mi) - 1.147E-9(Fi)$

گروه دهم $Di = 23.990 + 0.115(Li) - 0.079(Ri) + 0.378(Mi) - 1.390E-9(Fi)$

گروه یازدهم $Di = 19.111 - 0.109(Li) + 0.146(Ri) - 0.03(Mi) - 1.940E-9(Fi)$

گروه دوازدهم $Di = 20.198 - 0.051(Li) + 0.185(Ri) + 0.506(Mi) + 1.573E-8(Fi)$

گروه سیزدهم $Di = 29.009 + 0.108(Li) - 0.076(Ri) + 0.622(Mi) + 1.103E-8(Fi)$

گروه چهاردهم $Di = 17.746 + 0.10(Li) + 0.143(Ri) - 0.013(Mi) + 1.273E-8(Fi)$

گروه پانزدهم $Di = 21.729 + 0.101(Li) + 0.132(Ri) + 0.387(Mi) - 1.592E-9(Fi)$

گروه شانزدهم $Di = -27.054 + 0.104(Li) + 0.176(Ri) + 0.436(Mi) + 1.094E-8(Fi)$

بحث و نتیجه‌گیری

در این تحقیق اطلاعات ۲۲ شاخص‌های مرتبط با مشتریان وارد مدل‌های بیزین غیر خطی (BMA، TVP-DMA و TVP-DMS)، گردیدند. بر اساس میزان خطا مدل BMA از

بالاترین دقت برخوردار بود. پس از برآورد مدل ۴ متغیر اصلی شناسایی گردیدند. بر اساس نتایج متغیرهای موجودی حساب؛ مجموع مانده سپرده ها؛ مجموع مانده کل تسهیلات جاری و حجم تراکنش های مالی به عنوان متغیرهای غیر شکننده تعریف گردیدند.

در ادامه بر اساس نتایج مشخص گردید رویکرد C-MEANS نسبت به K-MEANS از دقت بالاتری برخوردار است و از میان رویکردهای فازی نوع ۱، نوع ۲ و نوع ۳ مدل فازی نوع سه در حالت ترکیب گوسی-مثلی از دقت بالاتری برخوردار است. بر اساس رویکرد C-MEANS، ۱۶ خوشه شناسایی گردیدند و ویژگی های هر ۱۶ خوشه مورد تجزیه و تحلیل قرار گرفت. بر اساس نتایج اکثر متغیرها تأثیر مثبت و معناداری بر وضعیت هر خوشه دارند.

توصیه های مبتنی بر مدل های بیزین

با توجه به اینکه متغیرهای متعددی بر خوشه بندی مشتریان موثر است و از طرفی میزان اهمیت آن ها در هر خوشه می تواند متفاوت باشد در سیستم توصیه گر طراحی شده برای هر خوشه باید متغیرهای اثرگذار و غیرشکننده را شناسایی نمود و بسته به متغیرهای غیر شکننده هر شاخص اقدام به توصیه نمود.

با توجه به اینکه در مدل های بیزین احتمال وقوع هر متغیر در طی زمان در حال تغییر است؛ بر این اساس وزن هر عامل در هر خوشه متغیر خواهد بود. لازم است در سیستم توصیه گری که طراحی می گردد پویای کافی جهت امکان تغییر مشتری در خوشه ها مورد بازبینی قرار گیرد.

توصیه های مبتنی بر خوشه بندی

با توجه به اینکه رویکردهای متفاوتی جهت خوشه بندی متغیرها مشتریان وجود دارد و هر یک از رویکردهای مورد بررسی تعداد خوشه بهینه را متفاوت ارائه می کند؛ استفاده از مدل های میانگین گیری پویا از خوشه های مشتریان می تواند بانک را در ارائه توصیه به خوشه های مختلف از گمراهی خارج نماید.

با توجه به اینکه در خوشه های مختلف شدت اثرگذاری و معناداری متفاوت است در

هر خوشه به تناسب میزان اثرگذاری و معناداری شاخص مربوط اقدام به ارائه توصیه سیاستی نمود.

توصیه‌های مبتنی بر فازی

با توجه به دقت بالاتر مدل‌های فازی نوع ۳ در تحقیق حاضر و مبتنی بر توزیع داده‌های گوسی و مثلثی لازم است با توجه به سطح ریسک‌پذیری مشتریان از توزیع‌های متناسب که توانایی پوشش دهی و شناسایی انواع ریسک از سمت مشتریان را داشته باشد؛ اقدام به ارائه توصیه نمود.

توصیه‌های مبتنی بر متغیرهای غیر شکننده

با توجه به تابعیت هر خوشه از متغیر مانده حساب و تأثیر مثبت این متغیر بر بهبود جایگاه هر خوشه در درآمدزایی بانک نیاز است، در سیستم توصیه‌گر پیشنهاداتی از قبیل استفاده از تسهیلات با بازپرداخت منقطع و وابسته به میزان خواب مانده حساب بلند مدت توجه نمود.

سپاسگزاری

نویسندگان از داوران ناشناس که در بهبود کیفیت مقاله کمک کردم اند تشکر می‌کنند.

تعارض منافع

تعارض منافع وجود ندارد.

پی‌نوشت‌ها

- | | |
|------------------------|--------------------------|
| 1. Kinge | 2. Vidal |
| 3. Theodoridis | 4. Zoynul Abedin |
| 5. Liu et al., | 6. Yuan et al., |
| 7. Aslam et al., | 8. Kalaivani and Sumathi |
| 9. De Caigny et al., | 10. Jain et al., |
| 11. Chen et al., | 12. Alam et al., |
| 13. Clerkin and Hanson | 14. Berggrun et al., |

- | | |
|-------------------------------------|----------------------------|
| 15. Amin et al. | 16. Naïve Bayes |
| 17. Bahnsen et al., | 18. Long et al., |
| 19. Abedin | 20. Zhang et al., |
| 21. Keramati et al., | 22. Mujica et al. |
| 23. Wojnarski | 24. Baumann et al. |
| 25. Share of Wallet | 26. Ngai et al. |
| 27. Fejza et al. | 28. Rahman and Khan |
| 29. University of California Irvine | 30. Kalaivani and Sumathi |
| 31. Decision Tree | 32. Multilayer Perceptron |
| 33. Logistic Regression | 34. Zhou et al. |
| 35. Raju and Dhandayudam | 36. Abbasimehr and Shabani |
| 37. De Caigny et al., | 38. Ho et al. |
| 39. Chen et al. | 40. Charte et al., |
| 41. Eigenvalues | 42. Canonical Correlation |

منابع

- خاتمی فیروزآبادی، محمدعلی، تقوی فرد، محمد تقی، سجادی، سید خلیل، بامدادصوفی، جهانیار. (۱۳۹۷). مدل بهینه‌سازی چندهدفه تخصیص خدمت به مشتریان بانک با به کارگیری داده کاوی و شبیه‌سازی. مطالعات مدیریت کسب و کار هوشمند، ۷(۲۵)، ۴۰-۵.
- سیفی، میثم، حیدری، زهرا، ابراهیمی، مریم. (۱۴۰۲). مدل‌سازی و شبیه‌سازی مدیریت الکترونیکی ارتباط با مشتریان با هدف مطالعه میزان وفاداری آن‌ها (مطالعه موردی: بانک تجارت). فصل‌نامه مطالعاتی در مدیریت بانکی و بانکداری اسلامی، ۸(پاییز)، ۱۳۹-۱۶۳.
- قبولی، ناصر، بافنده زنده، علیرضا، عالی، صمد. (۱۴۰۲). کشف دانش حاکم بر ویژگی‌های جمعیت‌شناختی مشتریان در انتخاب بانک‌ها با استفاده از قوانین انجمنی در داده کاوی. فصلنامه مهندسی مدیریت نوین، ۹(۳)، ۹۶-۱۲۱.
- Mohammad Zoynul Abedin, Petr Hajek, Taimur Sharif, Md. Shahriare Satu, Md. Imran Khan, (2023), Modelling bank customer behaviour using feature engineering and classification techniques, Research in International Business and Finance, Volume 65, April 2023, 101913
- Abbasimehr, H., Shabani, M., 2019. A new methodology for customer behavior analysis using time series clustering: A case study on a bank's customers. Kybernetes 50 (2), 221-242.
- Abedin, M.Z., Chi, G., Uddin, M.M., Satu, M.S., Khan, M.I., Hajek, P., 2020. Tax default prediction using feature transformation-based machine learning. IEEE Access 9, 19864-19881.

- Alam, N., Gao, J., Jones, S., 2021. Corporate failure prediction: An evaluation of deep learning vs discrete hazard models. *J. Int. Final. Inst. Money* 75, 101455. Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., Anwar, S., 2019. Customer churn prediction in telecommunication industry using data certainty. *J. Bus. Res.* 94, 290–301.
- Aslam, F., Hunjra, A.I., Ftiti, Z., Louhichi, W., Shams, T., 2022. Insurance fraud detection: Evidence from artificial intelligence and machine learning. *Res. Int. Bus. Finance* 62, 101744.
- Bahnsen, A.C., Aouada, D., Stojanovic, A., Ottersten, B., 2016. Feature engineering strategies for credit card fraud detection. *Expert Syst. Appl.* 51, 134–142.
- Baumann, C., Burton, S., Elliott, G., 2007. Predicting consumer behavior in retail banking. *J. Bus. Manag.* 13 (1), 79–96.
- Berggrun, L., Salamanca, J., Díaz, J., Ospina, J.D., 2020. Profitability and money propagation in communities of bank clients: A visual analytics approach. *Finance Res. Lett.* 37, 101387.
- Charte, D., Charte, F., Herrera, F., 2022. Reducing data complexity using autoencoders with class-informed loss functions. *IEEE Trans. Pattern Anal. Mach. Intell.* <http://dx.doi.org/10.1109/TPAMI.2021.3127698>.
- Chen, C., Geng, L., Zhou, S., 2021. Design and implementation of bank CRM system based on decision tree algorithm. *Neural Comput. Appl.* 33, 8237–8247.
- Clerkin, N., Hanson, A., 2021. Debit card incentives and consumer behavior: evidence using natural experiment methods. *J. Financ. Serv. Res.* 60 (2), 135–155.
- De Caigny, A., Coussement, K., De Bock, K.W., Lessmann, S., 2020. Incorporating textual information in customer churn prediction models based on a convolutional neural network. *Int. J. Forecast.* 36 (4), 1563–1578.
- Fejza, V., Livoreka, R., Bajrami, H., 2017. Analyzing consumer behavior in banking sector of Kosovo. *Eurasian J. Bus. Manag.* 5 (4), 33–48.
- Ho, S.C., Wong, K.C., Yau, Y.K., Yip, C.K., 2019. A machine learning approach for predicting bank customer behavior in the banking industry. In: *Machine Learning and Cognitive Science Applications in Cyber Security*. IGI Global, pp. 57–83.
- Jain, H., Yadav, G., Manoov, R., 2021. Churn prediction and retention in banking, telecom and IT sectors using machine learning techniques. In:

- Advances in Machine Learning and Computational Intelligence. Springer, Singapore, pp. 137–156.
- Kalaivani, D., Sumathi, P., 2019. Factor based prediction model for customer behavior analysis. *Int. J. Syst. Assur. Eng. Manag.* 10 (4), 519–524.
- Keramati, A., Ghaneei, H., Mirmohammadi, S.M., 2016. Developing a prediction model for customer churn from electronic banking services using data mining. *Financ. Innov.* 2 (1), 1–13.
- Kinge, A., Oswal, Y., Khangal, T., Kulkarni, N., Jha, P., 2022. Comparative study on different classification models for customer churn problem. In: *Machine Intelligence and Smart Systems*. Springer, Singapore, pp. 153–164.
- Liu, Y., Yang, M., Wang, Y., Li, Y., Xiong, T., Li, A., 2022. Applying machine learning algorithms to predict default probability in the online credit market: Evidence from China. *Int. Rev. Financ. Anal.* 79, 101971.
- Long, W., Lu, Z., Cui, L., 2019. Deep learning-based feature engineering for stock price movement prediction. *Knowl.-Based Syst.* 164, 163–173.
- Mujica, L.E., Melendez, J., Colomer, J., 2002. Modeling the bank's client behavior using case based reasoning and self-organizing map. (Accessed 20 December 2016).
- Ngai, E.W., Xiu, L., Chau, D.C., 2009. Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Syst. Appl.* 36 (2), 2592–2602.
- Rahman, A., Khan, M.N.A., 2018. A classification based model to assess customer behavior in banking sector. *Eng. Technol. Appl. Sci. Res.* 8 (3), 2949–2953.
- Raju, S.S., Dhandayudam, P., 2018. Prediction of customer behaviour analysis using classification algorithms. In: *AIP Conference Proceedings*, Vol. 1952 No. 1. AIP Publishing LLC, 020098.
- Theodoridis, G., Tsadiras, A., 2022. Applying machine learning techniques to predict and explain subscriber churn of an online drug information platform. *Neural Comput. Appl.* 1–14. <http://dx.doi.org/10.1007/s00521-022-07603-9>.
- Vidal, A., Kristjanpoller, W., 2020. Gold volatility prediction using a CNN-LSTM approach. *Expert Syst. Appl.* 157, 113481.
- Wojnarski, M., 2002. Modeling the Bank Client's Behavior with LTF-C Neural Network. Institute of Informatics, Warsaw University.

- Yuan, K., Chi, G., Zhou, Y., Yin, H., 2022. A novel two-stage hybrid default prediction model with k-means clustering and support vector domain description. *Res. Int. Bus. Finance* 59, 101536.
- Zhang, X., Han, Y., Xu, W., Wang, Q., 2021a. HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture. *Inform. Sci.* 557, 302–316.
- Zhang, H., Shi, Y., Yang, X., Zhou, R., 2021b. A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance. *Res. Int. Bus. Finance* 58, 101482.
- Zhou, X., Bargshady, G., Abdar, M., Tao, X., Gururajan, R., Chan, K.C., 2019. A case study of predicting banking customers behaviour by using data mining. In: 2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing. BESC, IEEE, pp. 1–6.

Designing a model for banking service allocation

Nazanin Kashani Kikoo¹

Mahnaz Rabiei²

Kiamars Fathi Hafshejani³

Abstract

Nowadays, the organizations', especially banks' interaction with customers has changed significantly as customer relationship management (CRM). The current research explains a model for customer relationship with a Fuzzy-Bayesian recommender approach. It is an applied and exploratory research, which has reviewed the information of 98,604 customers. It has used 3 approaches of data mining, fuzzy, non-linear Bayesian averaging. The information of 22 indicators of customers was inputted into non-linear Bayesian models (BMA, TVP-DMA and TVP-DMS). The BMA model, regarding its error rate, had the highest accuracy. As the results show, account balance, total balance of deposits, total balance of current facilities and volume of financial transactions were the research's non-fragile variables. C-MEANS approach is more accurate than K-MEANS. The C-MEANS approach could identify 16 clusters and analyze the characteristics of each of 16 clusters. Thus, most of the variables have a positive and significant effect on the status of each cluster.

Keywords: Gold banking services, customer relationship management, fuzzy clustering, data mining.

1. Ph.D. student, Department of Information Technology Management, South Tehran Branch, Islamic Azad University, Tehran, Iran

2. Department of Economics, Modeling and Optimization Research Center in Engineering Sciences, South Tehran Branch, Islamic Azad University, Tehran, Iran

3. Department of Industrial Management, South Tehran Branch, Islamic Azad University, Tehran, Iran