Journal of Information Technology Management
ISSN: 2008-5893
EISSN: 2423-5059
Vol. 9, No. 3; PP. 477-506
Fall 2017

University of Tehran
Faculty of Management

Designing a Predictive Analytics for the Formulation of Intelligent Decision Making Policies for VIP Customers Investing in the Bank

Iman Raeesi Vanani [1]

Abstract: Special, privileged or VIP customers are of great significance to the banks since they continuously and broadly invest in deposits and remain loyal to the banks. This loyalty is dependent on the broad and specific services they receive, deposit interests, and the tuned regulatory actions that banks take for according to the grade of special customers and their propensity to risk. In the current research, a dataset of two thousand ordinary and special privileged customers were collected according to their demographics, accounts information, and level of investment in the bank. The grade of special customer and their propensity to taking risks are also determined by the experts of the bank. Afterwards, a range of learning algorithms are applied for designing and validating classification and prediction methods on special customers' grades and their propensity to risk. Final results are then analyzed and prepared as a set of intelligent and improvable rules that assist the bank managers in formulating interactive and predictive decision making policies from the initiation of the customer relationship with the bank.

Key words: ***Bank investment, Classification, Intelligent decision making, Predictive analytics, Privileged customer.***

1. Assistant Prof. of Industrial Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran

Submitted: 06 / March / 2017
Accepted: 06 / August / 2017
Email: imanraeesi@atu.ac.ir

مدیریت فناوری اطلاعات
Journal of Information Technology Management
دانشکدهٔ مدیریت دانشگاه تهران
دورهٔ ۹، شمارهٔ ۳
پاییز ۱۳۹۶
صص. ۵۰۶- ۴۷۷

طراحی تحلیل‌های آینده‌نگر به‌منظور تدوین سیاست تصمیم‌گیری هوشمند برای مشتریان ویژهٔ سرمایه‌گذار در بانک

ایمان رئیسی وانانی[۱]

چکیده: مشتریان ویژه برای بانک‌ها اهمیت خاصی دارند، چراکه سرمایه‌گذاری مستمر در بانک انجام می‌دهند و معمولاً به یک بانک وفادارند. این وفاداری در گرو ارائهٔ خدمات خاص و متنوع به این مشتریان، ارائهٔ سود متناسب با سپرده و تعامل بانک بر اساس سطح مشتری ویژه و تمایل به ریسک‌پذیری آنهاست. در تحقیق حاضر، داده‌های دو هزار مشتری عادی و ویژه بر اساس ویژگی‌های جمعیت‌شناختی، حساب‌های بانکی و سرمایه‌گذاری در بانک گردآوری شده است. همچنین سطح مشتری ویژه و میزان تمایل به ریسک‌پذیری آنها نیز توسط بانک معین شده است؛ سپس، مجموعه‌ای از الگوریتم‌های یادگیرنده برای طبقه‌بندی و پیش‌بینی سطح مشتری ویژه و میزان تمایل آنها به ریسک توسط نرم‌افزار طراحی شده، اعتبارسنجی شده است. نتایج نهایی به‌صورت مجموعه‌ای از قواعد قابل بهبود و هوشمند تحلیل و آماده شده‌اند تا به مدیران بانک در تدوین سیاست‌های تصمیم‌گیری تعاملی و آینده‌نگر در ابتدای همکاری مشتری ویژه با بانک یاری رسانند.

واژه‌های کلیدی: ***تحلیل آینده‌نگر، تصمیم‌گیری هوشمند، سرمایه‌گذاری بانکی، طبقه‌بندی، مشتری ویژه.***

۱. استادیار مدیریت صنعتی، دانشکدهٔ مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۵/۱۲/۱۶
تاریخ پذیرش نهایی مقاله: ۱۳۹۶/۰۵/۱۵
E-mail: imanraeesi@atu.ac.ir

مقدمه

امروزه ساختن مدل‌ها و ابزارهای تصمیم‌گیری برای سازمان‌ها اهمیت بسزایی پیدا کرده است، سازمان‌ها تلاش می‌کنند با تخصیص منابع مناسب، مشتریان ارزشمند را حفظ کنند. این هدف با استفاده از مدیریت مشتری و ابزارهای پشتیبانی تصمیم‌گیری تحقق می‌یابد تا به سازمان‌ها کمک کند بین گروه‌های مختلف مشتریان تمایز قائل شوند. مشتریان وفادار، ارزش فراوانی برای سازمان‌ها ایجاد می‌کنند. آنها می‌توانند عمر طولانی و توانمندی رو به گسترش سازمان‌ها را تضمین کنند. این مشتریان، برای بانک‌ها نیز اهمیت بسیاری دارند. وفاداری مشتریان معمولاً از دو جنبهٔ وفاداری رفتاری و وفاداری نگرشی سنجیده می‌شود. هر دو جنبه، ناظر بر استمرار مشتریان در دریافت خدمات از بانک و نگرش آنها در انتقال این وفاداری به دیگران است (لی و پتریک، ۲۰۰۸). در واقع، وفاداری مشتری عبارت است از فراوانی یا کمیت دریافت محصولات یا خدمات از هر برند تجاری معتبر. این مشتریان به‌طور معمول به‌عنوان مشتریان ویژهٔ سازمان‌ها شناخته می‌شوند و رویکرد سازمان نسبت به آنها متفاوت خواهد بود.

کیفیت خدماتی که مشتریان دریافت می‌کنند، تابع متغیرهای مختلفی است. به‌طور کلی، کیفیت ادراک شدهٔ مشتری از خدمات دریافتی عبارت است از نتیجهٔ ارزیابی یا قضاوت مشتری دربارهٔ روش و کیفیتی که خدمات یا محصولات توسط تأمین‌کننده به مشتری ارائه شده است (انصاری و ریاسی، ۲۰۱۶، گالان، جارویس، براون و بیتنر، ۲۰۱۳). هر چند طی دههٔ گذشته، توسعهٔ فناوری‌های اطلاعاتی و ارتباطی در سطح صنعت مالی و بانکداری، تأثیر فراوانی بر نوع و کیفیت خدمات قابل ارائه به مشتریان داشته، این تأثیر بر مدل‌های کسب‌وکار بانک‌ها بیش از صنایع دیگر بوده است (کراکسیان و دانائیتا، ۲۰۱۴). یکی از فناوری‌های بسیار پرکاربرد که حوزه‌های علمی مختلف را دگرگون کرده است، علم داده و تحلیل‌های پیشرفته[1] است که شامل داده‌کاوی، متن‌کاوی، وب‌کاوی، جریان‌کاوی، تحلیل احساسات و سایر حوزه‌های ریاضی پیشرفته است (رضایی‌نور و شیخ بهایی، ۱۳۹۶؛ رئیسی وانانی و گنجعلی‌خان حاکمی، ۱۳۹۴). با استفاده از این الگوریتم‌ها، می‌توان روندها و نظم‌های پنهان در داده‌ها را به‌طور دقیق و عمیق تحلیل کرد و با آموزش دقیق و مستمر الگوریتم‌ها، امکان یادگیری از این روند و پیش‌بینی آینده را فراهم آورد (دیچو و دیچوا، ۲۰۱۷).

در تحقیق حاضر، مجموعه‌ای از داده‌های دو هزار مشتری عادی و ویژهٔ وفادار از یک بانک منتخب گردآوری شده است. این داده‌ها شامل ویژگی‌های جمعیت‌شناختی، مالی و روند سرمایه‌گذاری مشتریان ویژه است. در گام دوم، با استفاده از مجموعه‌ای از الگوریتم‌های

1. Data Science and Advanced Analytics

داده‌کاوی در علم داده، به شناسایی الگوهای منظم در داده‌ها و آموزش الگوریتم‌ها برای ایجاد قابلیت پیش‌بینی در آنها اقدام شده است. در نهایت، بر مبنای قابلیت‌های ایجاد شده در الگوریتم منتخب، امکان طراحی و پیاده‌سازی یک سیستم برای شناسایی هوشمند مشتریان ویژه در بدو ورود به بانک و تدوین سیاست‌های تعاملی مناسب به ازای مشتریان ویژه در آینده فراهم خواهد شد.

پیشینهٔ نظری پژوهش

هر بانکی برای تعریف مشتریان ویژه با توجه به نوع و دسته‌بندی و شناخت خود از مشتریان، شاخص متفاوتی دارد. به‌طور معمول، مشتریانی با درآمدهای بسیار زیاد که دارایی‌های کلان و ارزشمندی را در اختیار دارند، در بانک‌ها مشتریان ویژه محسوب می‌شوند (اسماعیل‌نیا، ۱۳۹۱). شرط این مهم، سرمایه‌گذاری مستمر این مشتریان در بانک و تمایل به ریسک و حفظ دارایی‌ها در بانک است. دو مورد از مهم‌ترین فعالیت‌های بانکی در تمام کشورها عبارت‌اند از (توتونچیان، ۱۳۷۵):

- قبول سپرده‌های پس‌انداز مدت‌دار و پرداخت بهرهٔ ماندهٔ این نوع حساب‌ها با توجه به مدت سپرده؛
- قبول سپردهٔ دیداری (حساب جاری) و پرداخت به مشتریان.

طبق تعریفی دیگر، سپرده‌های مردم نزد بانک‌ها به سپرده‌های دیداری[1] (قابل نقل و انتقال با چک)، سپرده‌های پس‌انداز[2] و سپرده‌های مدت‌دار[3] دسته‌بندی می‌شوند (منصف و منصوری، ۱۳۸۹).

سرمایه‌گذاری و ریسک‌پذیری مشتریان بانکی

به‌طورکلی، هرچه سرمایه‌گذاری ریسک بیشتری داشته باشد، پتانسیل پاداش و سود آن نیز بیشتر است؛ با این حال، ریسک ضرر بزرگ‌تری نیز در کمین بانک وجود دارد. هرچه ریسک کمتر باشد، بازده مورد انتظار نیز کمتر خواهد بود، اما ریسک زیان نیز کاهش می‌یابد. به‌عنوان بخشی از فرایند تخصیص دارایی‌ها که اولین و احتمالاً مهم‌ترین گام در سرمایه‌گذاری است، سطح تحمل ریسک یا تمایل به پذیرش ریسک فرد نیز تغییر خواهد کرد؛ به این صورت که چه میزان در دارایی‌های مختلف نظیر سهام، اوراق مشارکت، ارز، طلا، مسکن و بانک سرمایه‌گذاری

1. Demand Deposits
2. Saving Deposits
3. Time Deposits

می‌کند. تمایل به ریسک مشتری در بانک تابع چند متغیر است که به موقعیت خاص فرد بستگی دارد (دوموتور، ۲۰۱۱) از جمله:

- افق زمانی در سرمایه‌گذاری (بلندمدت، میان مدت یا کوتاه‌مدت)؛
- موقعیت شخصی فرد (مجرد، متأهل، جوان، میان‌سال، بازنشسته، سرپرست خانواده، دانشجو، تازه‌کار و...)؛
- سبک سرمایه‌گذاری فرد (تهاجمی، محافظه‌کارانه، با ریسک متوسط، ریسک‌گریز یا ریسک‌پذیر).

سرمایه‌گذاران الگوهای ریسک‌پذیری متفاوتی دارند که با توجه به میزان توانایی پذیرش ریسک یا احتمال موفق‌بودن، می‌توان آنها را به دو الگوی متفاوت دسته‌بندی کرد (دوموتور، ۲۰۱۳):

- ریسک‌پذیر: افرادی که از این الگو پیروی می‌کنند، در شرایط عدم اطمینان توانایی تحمل زیادی دارند و به همان میزان نیز خواهان دریافت بازده بیشتر هستند.
- ریسک‌گریز: کمابیش تمام افرادی که از این الگو پیروی می‌کنند، تحمل پذیرش هیچ‌گونه شرایط عدم اطمینان را ندارند و خواهان دریافت بازده کمتر، ولی تضمین شده‌اند.

از آنجا که ثروتمند ماندن دشوارتر از ثروتمند شدن است، مدیریت صحیح و منطقی ثروت و کنترل ریسک‌های متعددی که در بازار مالی وجود دارد، اهمیت بسیار زیادی دارد که این مسئله تنها از طریق برخورداری از دانش و آگاهی کامل نسبت به علم اقتصاد میسر می‌شود. بین دانشی که برای ثروتمند ماندن نیاز است و دانش ثروتمند شدن، تفاوت‌های بسیاری وجود دارد. بانکداری اختصاصی با مدیریت ثروت و کنترل خطرهای موجود می‌تواند ثروت موقت را به ثروت دائم تبدیل کند، بنابراین بررسی ابعاد مختلف بانکداری اختصاصی و فراهم کردن زیربناهای مورد نیاز آن در داخل کشور با توجه به پتانسیل موجود بسیار حائز اهمیت است (کاراسکو، پالاسیوس، کوادرادو، کرسپو و مزکوا، ۲۰۱۲).

پیشینهٔ تجربی

بانک‌ها تلاش می‌کنند با ارائهٔ خدمات بانکداری اختصاصی و پرهیز از ارائهٔ خدمات مشابه و مشترک به تمام مشتریان، رویکرد متفاوتی را نسبت به مشتریان ویژه اتخاذ کنند. بانکداری اختصاصی برای مشتریان ویژه، بر مواردی همچون ارائهٔ خدمات متمرکز، مشتریان معدود و خاص، امکانات منعطف، خدمات مالی و غیرمالی متنوع و مشتری‌محوری مستمر متمرکز می‌شود (غلام‌پورفرد، مختاری و رضوی خسروانی، ۱۳۹۱).

برخی از دانشمندان بر اهمیت احساسات و نگرش‌های مشتریان هنگام تصمیم‌گیری نسبت به استفاده از خدمات تأکید کرده و به سنجش آنها پرداخته‌اند (ریدت و کارسانا، ۲۰۱۷). در شرکت‌های خدماتی مانند بانک‌ها، نگرش احساسی نقش پررنگی در استفادهٔ مستمر از خدمات بانک دارد (استراکر و وریگلی، ۲۰۱۶). طراحی نگرش‌محور خدمات بانکی، یکی از رویکردهای نوینی است که بر اساس شناسایی به موقع و دقیق مشتریان ویژه، می‌تواند خدمات بانکی مناسب را به سمت این مشتریان سوق دهد (وریگلی، ۲۰۱۳). به حتم، تجربه‌های احساسی و نگرشی مثبت، به رفتار مثبت مشتریان ویژه نسبت به بانک و وفاداری آنها منجر خواهد شد (استراکر، وریگلی و روزمان، ۲۰۱۵).

با توجه به تجربه‌های موفق بانک‌های دنیا در ارائهٔ خدمات بانکداری اختصاصی به مشتریان ویژه، به نظر می‌رسد که این خدمت مزیت‌های کاربردی و متعددی برای بانک‌ها دارد که برخی از مهم‌ترین آنها را می‌توان به شرح زیر برشمرد (طاهری، ۱۳۹۱):

- ورود حجم زیادی از وجه اغنیای جامعه و خانواده‌های آنها با ماندگاری زیاد، به دلیل تمرکز حساب‌ها در بانک و به‌دنبال آن افزایش درآمدهای بهره‌ای؛
- افزایش درآمدهای غیربهره‌ای (کارمزدی) در نتیجهٔ ارائهٔ انواع متنوعی از خدمات مدیریت ثروت؛
- ایجاد حس وفاداری در مشتریان و خانواده‌های آنان به‌دلیل برقراری رابطهٔ بلندمدت با بانک؛
- فرصت‌جویی از فضای کمتر رقابتی بانکداری اختصاصی، به‌دلیل تخصصی بودن فعالیت‌های آن.

در این خصوص بانک‌ها با در نظر گرفتن وضعیت اقتصادی، اجتماعی، سیاسی و فرهنگی، برای سرمایه‌گذاری سودآور به مشتریان خود گزینه‌هایی نظیر موارد زیر را پیشنهاد می‌دهند که مشتریان بر اساس میزان تمایل به ریسک و مخاطره‌پذیری، در آنها سرمایه‌گذاری می‌کنند (ناظمی یگانه، ۱۳۸۹):

- سرمایه‌گذاری در انواع سپرده‌گذاری بورسی و پذیره‌نویسی؛
- سرمایه‌گذاری در صندوق‌های سرمایهٔ مشترک؛
- سرمایه‌گذاری در سهام شرکت‌های سهامی خاص؛
- سرمایه‌گذاری در صندوق‌های سرمایه‌گذاری در زمین و ساختمان
- سرمایه‌گذاری در ارزهای رایج.

این مطالعات، با پژوهش‌های گذشته نیز همخوانی دارند. برای نمونه، لیو (۲۰۰۱) با بررسی عملکرد پنج بانک فلیپینی که به بانکداری اختصاصی مشغول بودند، به این نتیجه دست یافت که این بانک‌ها همگام با رشد فناوری پیش می‌روند و دو موضوع مسئولیت اجتماعی و تعهد به جامعه برای آنها اهمیت ویژه‌ای دارد. همچنین استفادهٔ مؤثر از کمپین‌های مناسب تبلیغاتی، مسیر پیشرفت بانک‌ها را هموارتر کرده است.

شفیعی رودپشتی، حکاکی، جلالی و نوری (۱۳۹۳) مطالعه‌ای را با عنوان «تحلیلی بر وضعیت بانکداری اختصاصی در ایران» با هدف بررسی وضعیت بانکداری اختصاصی انجام دادند. با توجه به نتایج تحقیق، ۲۷ مورد از آسیب‌های موجود در راه توسعهٔ بانکداری اختصاصی شناسایی شد که در سه دسته آسیب‌های اجرایی ـ ساختاری، ادراکی ـ شناختی و آسیب‌های بدون خواست و اراده قرار گرفتند. شایان ذکر است که پنج عامل دولتی بودن بانک‌ها، کمبود تجهیزات و امکانات، کمبود منابع مالی، کمبود نیروی انسانی و نبودن سیستم جامع استعدادیابی منابع انسانی و بی‌ثباتی اقتصادی، به‌عنوان آسیب‌های اصلی شناخته شدند.

در پژوهش حاضر به‌منظور رفع مسائل ذکر شده و حرکت به سمت هوشمندسازی تصمیم‌گیری دربارهٔ مشتریان ویژه در بانک‌ها، از الگوریتم‌های داده‌کاوی استفاده شده است. در ادامه به‌طور مختصر به توصیف داده‌کاوی پرداخته شده است.

داده‌کاوی و فرایند آن

تعاریف مختلفی از داده‌کاوی ارائه شده است. داده‌کاوی به بررسی و تجزیه و تحلیل مقادیر عظیمی از داده‌ها به‌منظور کشف الگوی قوانین پنهان و معنادار درون داده‌ها گفته می‌شود (شهرابی و شجاعی، ۱۳۹۰). در تعریف دیگر، داده‌کاوی به فرایند نیمه‌خودکار تجزیه و تحلیل پایگاه داده‌های بزرگ به‌منظور یافتن الگوهای مفید اطلاق شده است (هان، کمبر و پی، ۲۰۱۱). ویتن و فرانک (۲۰۰۵) نیز، داده‌کاوی را فرایند استخراج اطلاعات معتبر، از پیش ناشناخته، فهم‌پذیر و قابل اعتماد از پایگاه داده‌های بزرگ و استفاده از آن در تصمیم‌گیری برای پیاده‌سازی فعالیت‌های تجاری مهم تعریف کرده‌اند.

داده‌کاوی فرایند کشف دانش پنهان درون داده‌هاست که با توصیف، تشریح، پیش‌بینی و کنترل پدیده‌های گوناگون، کاربرد بسیار وسیعی در حوزه‌های مختلف دارد؛ به‌گونه‌ای که مرز و محدودیتی برای آن در نظر گرفته نشده و از ذرات کف اقیانوس تا اعماق فضا کاربرد دارد. داده‌کاوی نوعی جنبش میان‌رشته‌ای است که حوزه‌هایی از قبیل پایگاه‌های اطلاعاتی، آمار، یادگیری ماشینی، محاسبات سریع، مصورسازی و ریاضیات را دربرمی‌گیرد (مارکو و راسل، ۲۰۰۹).

به‌طور کلی داده‌کاوی دارای دو رویکرد یادگیری نظارت‌شده[1] و بدون نظارت[2] است (شهرابی و شجاعی، ۱۳۹۰). در داده‌کاوی نظارت شده، با داشتن یک متغیر هدف خاص و از پیش تعیین شده، الگوریتم به‌دنبال الگوی خاص خواهد بود، در حالی که در داده‌کاوی بدون نظارت، متغیر هدفی وجود ندارد و هدف یافتن تشابهات بین گروه‌های همانند از اطلاعات است. روش‌های داده‌کاوی را براساس فعالیت و وظیفه به گروه‌های زیر می‌توان طبقه‌بندی کرد (لاروس، ۲۰۱۴):

۱. دسته‌بندی (طبقه‌بندی)
۲. تخمین
۳. پیش‌بینی
۴. همبستگی
۵. خوشه‌بندی
۶. توصیف و نمایه‌سازی

اولین مدل فرایندی داده‌کاوی توسط فیاد و همکارانش ارائه شده است (هان و همکاران، ۲۰۱۱). این مدل شامل نه فاز است: فاز اول با هدف شناسایی حوزهٔ کاربرد و درک قلمرو اجرا شده و هدف فرایند داده‌کاوی از دید کاربران شناسایی می‌شود. در فاز دوم، داده‌های مدنظر انتخاب و ایجاد می‌شوند تا فرایند کشف دانش روی آنها صورت گیرد. در فاز سوم، پالایش[3] و پیش‌پردازش داده‌ها انجام می‌شود که شامل فعالیت‌هایی مانند حذف موارد دارای اختلال و پرکردن مقادیر داده‌ای است که هنگام ورود اطلاعات در دسترس نبوده‌اند. در گام چهارم به تبدیل و کاهش ابعاد داده‌ها اقدام می‌شود. گام پنجم به انتخاب روش داده‌کاوی مدنظر، مانند طبقه‌بندی، خوشه‌بندی و... اختصاص دارد. فاز ششم، شامل پیاده‌سازی الگوریتم داده‌کاوی مدنظر و تعیین و تنظیم پارامترهای آن است. در گام هفتم، الگوریتم انتخاب شده در مرحلهٔ قبل روی داده‌ها اعمال می‌شود. گام هشتم تفسیر نتایج حاصل از اجرای الگوریتم است و در نهایت، در گام نهم دانش کشف شده از داده‌ها استفاده و مستندسازی شده و در سیستم‌های موجود کسب‌وکار به بهره‌برداری می‌رسد. فرایند داده‌کاوی و کشف دانش، نوعی فرایند تعاملی بین کاربران و کاوشگران داده است و به حلقه‌های تکرار بین مراحل مختلف نیاز دارد.

پیش‌بینی و طبقه‌بندی

عبارت پیش‌بینی به هر دو حالت پیش‌بینی مقادیر عددی و پیش‌بینی کلاس‌ها اشاره دارد. پیش‌بینی شامل شناسایی روند توزیع براساس اطلاعات در دسترس است. طبقه‌بندی و پیش‌بینی

1. Supervised
2. Unsupervised
3. Cleansing

ممکن است به تجزیه و تحلیل داده‌ها و متون علمی مرتبط نیاز داشته باشد تا صفاتی را که در فرایند طبقه‌بندی یا پیش‌بینی مشارکت و اثرگذاری ندارند، شناسایی شده و در صورت تمایل حذف شوند (هان و کمبر، ۲۰۰۶). الگوریتم‌های مختلفی که مبتنی بر شبکه‌های عصبی، ماشین‌های بردار پشتیبان، توابع شعاعی پایه، رگرسیون و سایر روش‌های پیوسته‌اند، قادر به پیش‌بینی دقیق با بیشترین جزئیات هستند.

از سوی دیگر، طبقه‌بندی فرایند جست‌وجوی مجموعه‌ای از مدل‌ها یا توابع است که کلاس‌های داده یا مفاهیم را توصیف یا تفکیک می‌کند تا بتوان از این مدل‌ها برای پیش‌بینی کلاس اشیا استفاده کرد. مدل استخراج شده، بر پایهٔ تحلیل مجموعه‌ای از داده‌های آموزشی یا داده‌هایی است که کلاس و طبقهٔ آنها از پیش مشخص است. روش‌هایی مانند درخت تصمیم، k نزدیک‌ترین همسایه، الگوریتم بیزین ساده و کمی‌سازی بردار یادگیرنده، در حوزهٔ الگوریتم‌های طبقه‌بندی قرار می‌گیرند (هان و کمبر، ۲۰۰۶). برای کسب اطلاعات کامل درباره الگوریتم‌های داده‌کاوی و انواع طبقه‌بندی آنها، مقالات متعددی وجود دارند که یکی از جامع‌ترین آنها، مقالهٔ لیائو، چو و هسیائو (۲۰۱۲) است.

روش‌شناسی پژوهش

تحقیق حاضر بر اساس هدف از نوع کاربردی و به لحاظ گردآوری داده‌ها از نوع کمی است. تحقیق کاربردی در جست‌وجوی دستیابی به هدف عملی است (دلاور، ۱۳۹۲) به بیان دیگر تحقیقات کاربردی به سمت کاربرد عملی دانش هدایت می‌شوند (بازرگان و همکاران، ۱۳۸۶). تحقیق حاضر برمبنای اطلاعات جمعیت‌شناختی، تراکنش‌های مالی و سرمایه‌گذاری‌های سپرده‌محور دو هزار نفر از مشتریان عادی و ویژهٔ یکی از بانک‌های بزرگ کشور در تهران و شهرستان است که به‌صورت کاملاً تصادفی انتخاب شده‌اند. اطلاعات جمع‌آوری شده مربوط به مشتریانی است که در آخرین ماه گردآوری اطلاعات در وضعیت فعال قرار داشتند.

در این تحقیق از گام‌های روش‌شناسی CRISP-DM که یکی از روش‌های استاندارد و پذیرفته‌شده برای انجام پروژه‌های داده‌کاوی است، استفاده می‌شود. این روش‌شناسی، یک مدل فرایندی برای داده‌کاوی ارائه می‌دهد که چرخهٔ عمر هر پروژهٔ داده‌کاوی را مرور می‌کند و شامل مراحل متناظر با یک پروژه، وظایف مربوطه و ارتباط بین این وظایف است (هان و کمبر، ۲۰۰۶؛ لاروس، ۲۰۱۴). این مراحل که در تحقیق حاضر نیز مدنظر قرار گرفته‌اند، به‌طور خلاصه عبارت‌اند از:

۱. شناخت کسب‌وکار[1]: در این مرحله به بررسی و شناخت کسب‌وکار مدنظر پرداخته می‌شود تا اهداف و عوامل موفقیت آنها مشخص شود. پس از تدوین اهداف پروژه، وضعیت موجود سازمان شناخته خواهد شد و در نهایت منابع، نیازها و محدودیت‌های موجود تعیین می‌شوند.

۲. شناخت داده‌ها[2]: این مرحله با جمع‌آوری داده‌ها آغاز می‌شود و با بررسی کیفیت داده‌ها، آشنایی با ساختار داده‌ها و شکل‌دهی به فرضیه‌هایی در خصوص داده‌های موجود پیش می‌رود.

۳. آماده‌سازی داده‌ها[3]: در این گام داده‌های مد نظر انتخاب خواهند شد؛ سپس روی این داده‌ها عملیات پالایش، آماده‌سازی، تبدیل و یکپارچه‌سازی انجام می‌شود تا از رسیدن به داده‌های با کیفیت، اطمینان لازم حاصل شود.

۴. مدل‌سازی[4]: در این مرحله با استفاده از روش‌های مدل‌سازی داده‌کاوی، یک مدل نمونه ساخته می‌شود و پس از هر مرحلهٔ اجرا به‌دقت اعتبارسنجی[5] می‌گردند تا کیفیت مدل ایجاد شده سنجیده شود. اگر مدل انتخاب شده کیفیت مدنظر را نداشته باشد، باید پارامترهای مدل را تغییر داد و مدل را بار دیگر اعتبارسنجی کرد تا در نهایت یک مدل با کیفیت به‌دست آید.

۵. ارزیابی[6]: پس از مدل‌سازی، به ارزیابی نتایج حاصل از مدل پرداخته می‌شود و متناسب با مدل ساخته‌شده، اعتبار مدل نهایی در محیط کار واقعی بررسی خواهد شد.

۶. پیاده‌سازی و توسعه[7] : در این گام، طرح توسعه ارائه می‌شود. همچنین طرح نظارت و نگهداری پس از پایان پروژه نیز در همین گام تهیه می‌شود. سرانجام یک گزارش نهایی و کامل از پروژه، استخراج شده و به مدیریت تقدیم خواهد شد.

همچنین از نرم افزارهای MATLAB R2017a و Rapidminer 7.5 برای انجام تحلیل‌ها بهره برده شده است. متغیرهایی که در پژوهش حاضر در دسترس بودند و از طریق سرورهای بانک استخراج شدند، به شرح جدول ۱ است. این متغیرها توسط خبرگان بانکی بررسی و تأیید شدند و داده‌های مربوط به هر یک از آنها قابلیت اتکا دارند.

---

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Validation
6. Evaluation
7. Deployment

جدول ۱. متغیرهای استفاده‌شده در تحلیل داده

ردیف	نوع متغیر	نام متغیر	محتوای داده متغیر
۱	جمعیت‌شناختی	جنسیت	مرد / زن
۲		سن	متغیر عددی
۳		تحصیلات	زیر دیپلم / دیپلم / کارشناسی / کارشناسی ارشد / دکتری
۴		وضعیت تأهل	مجرد / متأهل
۵		شغل	آزاد / کارمند (بخش دولتی یا خصوصی)
۶	فعالیت مالی	معدل موجودی	متغیر عددی
۷		گردش بستانکار	متغیر عددی
۸		سابقهٔ معوقات	بلی / خیر
۹		سابقهٔ چک برگشتی	بلی / خیر
۱۰		پروفایل اعتباری در سیستم بانکی	مطلوب / نامطلوب
۱۱	متغیر هدف	مشتری ویژه	بلی / خیر
۱۲	متغیر هدف	رتبهٔ مشتری ویژه	طلایی / نقره‌ای / پلاتینیوم
۱۳	سرمایه‌گذاری بانکی	زمینهٔ سرمایه‌گذاری	طلا، فارکس، سهام، مسکن / سپرده‌گذاری
۱۴		مدت سرمایه‌گذاری	کوتاه‌مدت / بلندمدت
۱۵		سابقهٔ زیان مالی	بلی / خیر
۱۶		هدف از سرمایه‌گذاری	حفظ ارزش سرمایه/ ایجاد جریان مالی ثابت/ افزایش سرمایه
۱۷		میزان سرمایه‌گذاری	کمتر از یک چهارم سرمایه / بین یک چهارم تا نیمی از سرمایه / بیش از نیمی از سرمایه
۱۸	متغیر هدف	میزان ریسک‌پذیری	محتاط / متعادل / مخاطره‌جو

سؤال‌های پژوهش

بر اساس متغیرهای در دسترس در داده‌های گردآوری شده، مهم‌ترین سؤال‌هایی که به تصمیم‌گیری مدیران نسبت به مشتری ویژه منجر می‌شوند، به شرح زیر تعریف شده‌اند. پاسخ به این سؤال‌ها توسط الگوریتم‌هایی که در بخش‌های بعدی استفاده می‌شوند، ارائه خواهد شد.

- کدام متغیرها بیشترین توانایی را در شناسایی رفتار و پیش‌بینی تبدیل مشتری به مشتری ویژه در بدو آغاز به همکاری دارند؟

- آیا امکان پیش‌بینی رتبهٔ (سطح) مشتریان ویژه، پیش از آغاز به همکاری با آنها وجود دارد؟
- آیا متغیرهای ارائه شده، می‌توانند میزان تمایل به ریسک سرمایه‌گذاری مشتریان ویژه را پیش‌بینی کنند؟

پس از شناسایی متغیرها، سؤال‌های پژوهش و داده‌های گردآوری شده، ابتدا داده‌ها پیش‌پردازش و پالایش می‌شوند، سپس الگوریتم‌ها اجرا و اعتبارسنجی می‌گردند و در نهایت، یافته‌ها تحلیل و بررسی خواهند شد.

پیش‌پردازش داده‌ها

در داده‌کاوی، همواره اصل بر حفظ حداکثری داده‌ها و فیلدهای اطلاعاتی و پرهیز از حذف آنها به‌دلیل ارزش زیاد اطلاعات به‌دست آمده است. داده‌های پژوهش حاضر نیز از این اصل مستثنی نیستند و پس از دقت و پیگیری بسیار، از طریق واحد فناوری اطلاعات و با استفاده از انواع روش‌های پالایش داده‌ها در پایگاه داده به‌دست آمده‌اند. به همین دلیل تلاش می‌شود داده‌ها حفظ شوند، مگر دلیل محکمی برای حذف برخی از رکوردها شناسایی شود.

شناسایی مقادیر خارج از محدوده

بر مبنای متغیرهایی که به شناسایی مشتریان ویژه منجر شده‌اند، به تحلیل مقادیر خارج از محدوده اقدام می‌شود. برای این منظور، از روش‌های عامل خارج از محدودهٔ محلی (LOF)[1] و شناسایی مقادیر خارج از محدوده بر اساس فاصله[2] استفاده شده است. با توجه به اینکه روش مبتنی بر فاصله، تعداد مشخص شدهٔ مقادیر خارج از محدوده توسط محقق را برای شناسایی مقادیر در نظر می‌گیرد، باید از روش دیگری برای بررسی بیشتر مقادیر خارج از محدوده نیز به‌طور موازی کمک گرفته شود. به این منظور با استفاده از روش LOF، تعداد تقریبی مقادیر مطابق با شکل ۱ به‌دست آمد که برابر با ۱۰ رکورد دادهٔ خارج از محدوده است. در حقیقت، ۱۰ رکورد اول در مقدار عدد خارج از محدوده، فاصلهٔ محسوسی با سایر رکوردهای داده دارند. نتیجهٔ حاصل از روش مبتنی بر فاصله نیز شباهت زیادی به نتایج روش LOF دارد. علاوه بر رکوردهای مشابه بین دو روش، چند رکورد دیگر نیز توسط این روش مشخص شد که آنها نیز در تحلیل‌ها مشارکت داده نمی‌شوند (شکل ۲).

---

1. Local Outlier Factor
2. Distance-Based Outlier Detection

ExampleSet (2000 examples, 1 special attribute, 10 regular attributes) Filter (2,000 / 2,000 examples):

Row No.	outlier ↓	Sex	Age	Education	MaritalStatus	Occupation	AVGBalance	CreditTurnO...	DueClaims	BouncedCh...	CreditEvalua...
1642	44.769	1	54	1	2	2	954386310	324090577	0	0	1
1977	31.986	1	35	2	2	1	4597748196	1658474581	0	0	1
1887	29.129	1	39	2	2	1	1590398672	490456004	0	0	1
1443	15.411	1	58	1	2	1	8490574404	280029444	0	0	1
1449	13.028	1	43	2	2	1	19705396455	633869446	0	0	1
1445	10.673	1	54	1	2	1	9671006477	203505870	0	0	1
1455	7.757	1	38	1	2	1	17975006752	721368533	0	0	1
1526	7.624	1	42	2	2	1	1048786005	42964733	0	0	1
1451	7.150	1	57	2	2	1	17485530285	324785234	0	0	1
1965	6.113	1	43	2	2	1	2756252108	698836539	0	0	1
1966	5.974	1	43	2	2	1	2786753508	689521039	0	0	1
1457	5.829	1	64	1	2	1	876958327	35674202	0	0	1
1726	5.706	1	43	2	2	1	2716451507	698351039	0	0	1
1485	5.601	1	43	2	2	1	2706652508	698752039	0	0	1
1579	5.499	1	43	3	2	1	3558470351	145084105	0	0	1
1606	5.439	1	43	2	2	1	2706652508	674591039	0	0	1

شکل ۱. مقادیر خارج از محدوده بر مبنای روش LOF برای متغیرهای جمعیت‌شناختی و فعالیت مالی

ExampleSet (2000 examples, 1 special attribute, 10 regular attributes) Filter (2,000 / 2,000 examples):

Row No.	outlier ↓	Sex	Age	Education	MaritalStatus	Occupation	AVGBalance	CreditTurnO...	DueClaims	BouncedCh...	CreditEvalua...
1443	true	1	58	1	2	1	8490574404	280029444	0	0	1
1449	true	1	43	2	2	1	19705396455	633869446	0	0	1
1451	true	1	57	2	2	1	17485530285	324785234	0	0	1
1455	true	1	38	1	2	1	17975006752	721368533	0	0	1
1485	true	1	43	2	2	1	2706652508	698752039	0	0	1
1565	true	1	54	1	2	1	8126916477	225785870	0	0	1
1696	true	1	38	1	2	1	15871005752	774360533	0	0	1
1725	true	1	43	2	2	1	2405651508	698865039	0	0	1
1936	true	1	38	1	2	1	15980201732	775260533	0	0	1
1977	true	1	35	2	2	1	4597748196	1658474581	0	0	1
1	false	1	68	2	2	2	21480300	3908146	0	1	1
2	false	2	49	2	1	2	43602741	3103785	0	0	1
3	false	2	40	2	1	2	901202271	3617510	0	0	1
4	false	1	32	2	1	2	13589545	3972440	0	0	1
5	false	1	35	3	2	1	500679203	15898623	0	0	1
6	false	1	31	1	1	2	83597471	5623951	0	0	1

شکل ۲. مقادیر خارج از محدوده بر مبنای روش مبتنی بر فاصله برای متغیرهای جمعیت‌شناختی و فعالیت مالی

در گام دوم، مقادیر خارج از محدودهٔ مشتریان ویژه بر اساس متغیرهای مربوط به سرمایه‌گذاری در بانک سنجیده شدند و مشخص شد که تعداد رکوردهای خارج از محدوده بسیار کم هستند و تأثیری در نتیجهٔ نهایی ندارند. نتایج تحلیل در شکل ۳ ارائه شده است. پس از شناسایی این مقادیر، با توجه به تعداد کم رکوردهای خارج از محدوده و رکوردهای دارای اختلال[1]، این امکان وجود دارد که بتوان الگوریتم‌های داده‌کاوی را با دقت مناسبی آموزش داد و برای پیش‌بینی استفاده کرد. به همین منظور، در گام بعد بدون در نظر گرفتن این مقادیر، الگوریتم‌ها روی کلیهٔ داده‌ها اعمال شدند تا میزان دقت پیش‌بینی‌ها محاسبه شود. پس از

1. Noise

پیاده‌سازی الگوریتم‌ها و اعتبارسنجی آنها، مشخص شد که مقادیر خارج از محدوده را می‌توان با دقت بسیار مناسبی پیش‌بینی کرد. از آنجا که تعداد مشتریان ویژه در هر بانک بسیار محدودند، تصمیم بر آن شد که مشتریان ویژهٔ خارج از محدوده و دارای شرایط خاص نیز، به‌طور کامل توسط الگوریتم‌ها لحاظ شده و تحلیل‌ها انجام شوند تا از تمام پتانسیل موجود در داده استفاده شود؛ مگر آن که در نتایج الگوریتم‌ها انحراف غیرقابل رفعی ایجاد شود که در نهایت دقت نتایج در سطح مطلوبی ارزیابی شد.

ExampleSet (600 examples, 1 special attribute, 5 regular attributes)

Row No.	outlier ↓	InvestmentT...	InvestmentD...	LossExperie...	InvestmentG...	InvestmentA...
33	0.314	1	2	0	3	2
34	0.314	1	2	0	3	2
93	0.314	1	2	0	3	2
94	0.314	1	2	0	3	2
153	0.314	1	2	0	3	2
154	0.314	1	2	0	3	2
213	0.314	1	2	0	3	2
214	0.314	1	2	0	3	2
273	0.314	1	2	0	3	2
274	0.314	1	2	0	3	2
333	0.314	1	2	0	3	2

ExampleSet (600 examples, 1 special attribute, 5 regular attributes)

Row No.	outlier	InvestmentT...	InvestmentD...	LossExperie...	InvestmentG...	InvestmentA...
1	true	2	1	0	2	0
2	true	2	1	0	2	0
3	true	2	1	0	2	0
4	true	2	1	0	2	0
5	true	1	2	1	3	1
6	false	1	2	1	3	1
7	false	1	2	0	3	1
8	false	1	2	0	3	1
9	false	1	2	1	3	1
10	false	1	2	1	3	1
11	false	1	2	0	3	1

شکل ۳. مقادیر خارج از محدوده بر مبنای روش مبتنی بر فاصله و LOF برای متغیرهای سرمایه‌گذاری بانکی

تحلیل همبستگی داده‌ها

پس از تحلیل اولیهٔ مقادیر خارج از محدوده، به تحلیل همبستگی داده‌ها اقدام شد. بر اساس این که اغلب فیلدهای اطلاعاتی در حالت دو یا چند وضعیتی قرار دارند، تحلیل همبستگی آنها دقیق و قابل اتکا نخواهد بود. از سوی دیگر، امکان ورود آنها به الگوریتم‌های متناسب با فیلدهای چند حالتی مهیا است. با وجود این، به‌منظور تحلیل دقیق‌تر داده‌ها، همبستگی تقریبی آنها سنجیده شد که نتیجهٔ آن در شکل ۴ ارائه شده است. تنها فیلدهای گردش مالی و ماندهٔ حساب، همبستگی نسبتاً زیادی دارند که به‌دلیل ارتباط حجم تراکنش‌ها با ماندهٔ حساب است. این دو فیلد به‌دلیل ماهیت متفاوت، باید در تحلیل‌ها در نظر گرفته شوند. به همین دلیل نیازی به حذف یکی از دو فیلد نیست. نتایج نهایی الگوریتم‌ها نیز نشان‌دهندهٔ دقت زیاد یافته‌ها است. با توجه به اینکه بقیهٔ فیلدها همبستگی چندانی ندارند، می‌توان از آنها به‌طور مستقیم در الگوریتم‌های پیش‌بینی‌کننده استفاده کرد.

Attributes	Sex	Age	Education	MaritalStatus	Occupation	AVGBalance	CreditTurnOver	DueClaims	BouncedCheck	CreditEvaluation
Sex	1	-0.186	0.235	-0.100	0.243	-0.296	-0.250	-0.050	0.072	-0.079
Age	-0.186	1	-0.443	0.465	-0.134	0.174	0.110	-0.204	0.017	0.173
Education	0.235	-0.443	1	-0.335	0.107	-0.158	-0.111	-0.013	0.047	-0.005
MaritalStatus	-0.100	0.465	-0.335	1	-0.320	0.259	0.221	-0.263	-0.091	0.204
Occupation	0.243	-0.134	0.107	-0.320	1	-0.334	-0.293	0.029	0.214	-0.300
AVGBalance	-0.296	0.174	-0.158	0.259	-0.334	1	0.879	-0.139	-0.122	0.200
CreditTurnOver	-0.250	0.110	-0.111	0.221	-0.293	0.879	1	-0.113	-0.098	0.164
DueClaims	-0.050	-0.204	-0.013	-0.263	0.029	-0.139	-0.113	1	0.012	-0.253
BouncedCheck	0.072	0.017	0.047	-0.091	0.214	-0.122	-0.098	0.012	1	-0.308
CreditEvaluation	-0.079	0.173	-0.005	0.204	-0.300	0.200	0.164	-0.253	-0.308	1

شکل ٤. نتایج تحلیل همبستگی متغیرهای تحقیق

در ادامهٔ پیش‌پردازش داده‌ها و قبل از انجام داده‌کاوی، از الگوریتم‌های تحلیل مؤلفه‌های اصلی[1] و تجزیهٔ مقدار منفرد[2] نیز استفاده شد که چون تأثیر معنادار و مهمی بر نتایج داده‌کاوی نداشتند، نتایج آنها در مقاله ذکر نشد. در نهایت، کلیهٔ متغیرها و رکوردها مدنظر قرار گرفتند و به‌دلیل نتایج مطلوب الگوریتم‌های داده‌کاوی، هیچ‌یک از فیلدها یا رکوردها حذف نشدند.

پیاده‌سازی الگوریتم‌ها و تحلیل یافته‌ها

برای پیاده‌سازی طبقه‌بندی، الگوریتم‌های بسیاری آزمون شدند که از بین آنها، الگوریتم‌های درخت تصمیم؛ K نزدیک‌ترین همسایه؛ شبکه‌های بزیین[3] و شبکه‌های عصبی چندلایه[4] بهترین

1. Principal Component Analysis (PCA)
2. Singular Value Decomposition (SVD)
3. Naïve Bayes
4. Neural Networks

نتایج بودند و به‌عنوان الگوریتم‌های منتخب، برای طراحی، پیاده‌سازی، اعتبارسنجی و تحلیل یافته‌ها استفاده شدند.

طبقه‌بندی و پیش بینی مشتری ویژه

در گام نخست داده‌کاوی، بر اساس فیلدهای ورودی‌ای که بر مشتری ویژه اثرگذارند و در بخش قبل ذکر شده بودند، به آموزش مدل طبقه‌بندی اقدام شد. نتایج پیاده‌سازی و اعتبارسنجی الگوریتم‌ها (شکل ۵) و همچنین درخت تصمیم به‌دست آمده، نشان‌دهندهٔ این مهم است که با استفاده از متغیرهای ماندهٔ حساب و گردش بستانکار مالی، با دقت بسیار زیادی می‌توان مشتریان ویژه را شناسایی کرد که با رویکرد مدیریتی و تصمیم‌گیری بانک کاملاً تطابق دارد.

دقت الگوریتم K-NN بیش از سایر الگوریتم‌ها و برابر با ۱۰۰ درصد به‌دست آمد، اما بیش‌برازش زیادی روی داده‌ها ایجاد کرده است؛ در صورتی که داده‌های مشابهی در اختیار این الگوریتم قرار گیرد، اصولاً دقت تخمین کمتری ارائه می‌کند. دقت الگوریتم‌های درخت تصمیم و شبکه‌های عصبی بسیار مناسب هستند، فقط دقت روش بیزین به‌طور نسبی کمتر است.

accuracy: 99.95%

	true No	true Yes	class precision
pred. No	1399	0	100.00%
pred. Yes	1	600	99.83%
class recall	99.93%	100.00%	

accuracy: 99.95%

	true No	true Yes	class precision
pred. No	1399	0	100.00%
pred. Yes	1	600	99.83%
class recall	99.93%	100.00%	

accuracy: 73.20%

	true No	true Yes	class precision
pred. No	864	0	100.00%
pred. Yes	536	600	52.82%
class recall	61.71%	100.00%	

accuracy: 100.00%

	true No	true Yes	class precision
pred. No	1400	0	100.00%
pred. Yes	0	600	100.00%
class recall	100.00%	100.00%	

شکل ۵. دقت الگوریتم‌های طبقه‌بندی

(به‌ترتیب: شبکه‌های عصبی، درخت تصمیم، شبکه‌های بیزین و K نزدیک‌ترین همسایه)

با وجود دستیابی به نتایج بالا، هنوز به سؤال اصلی پژوهش (آیا می‌توان بدون توجه به گردش بستانکار مشتری و ماندهٔ حساب، به طبقه‌بندی مشتریان ویژه و شناسایی آنها از ابتدای ورود به بانک اقدام کرد؟) پاسخ داده نشده است. برای پاسخ به این سؤال، فیلدهای مذکور حذف شدند و بار دیگر طبقه‌بندی داده‌ها انجام گرفت. بر پایهٔ نتایج به‌دست آمده، مشاهده می‌شود که همچنان الگوریتم‌ها از دقت بسیار بالایی برخوردارند و بدون توجه به گردش و ماندهٔ حساب نیز می‌توان مشتریان ویژه را شناسایی کرد (شکل ۶).

accuracy: 93.00%

	true No	true Yes	class precision
pred. No	1280	20	98.46%
pred. Yes	120	580	82.86%
class recall	91.43%	96.67%	

accuracy: 96.10%

	true No	true Yes	class precision
pred. No	1382	60	95.84%
pred. Yes	18	540	96.77%
class recall	98.71%	90.00%	

accuracy: 70.90%

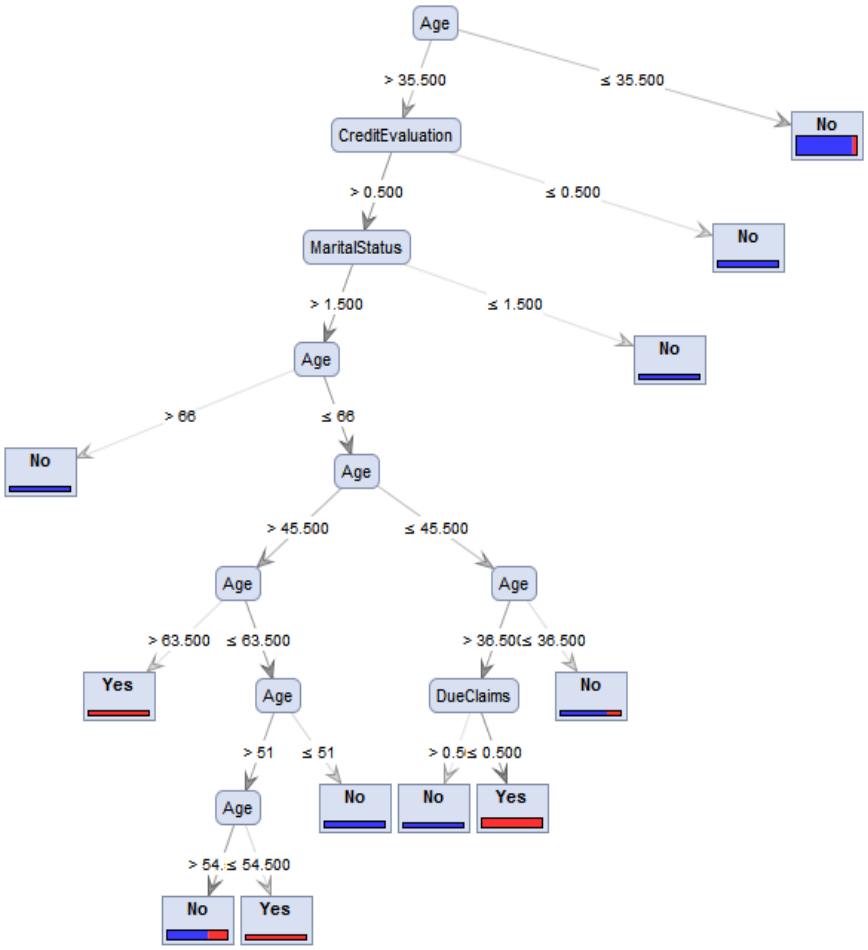
	true No	true Yes	class precision
pred. No	818	0	100.00%
pred. Yes	582	600	50.76%
class recall	58.43%	100.00%	

accuracy: 96.50%

	true No	true Yes	class precision
pred. No	1400	70	95.24%
pred. Yes	0	530	100.00%
class recall	100.00%	88.33%	

شکل ٦. دقت الگوریتم‌های طبقه‌بندی پس از حذف فیلدهای ماندهٔ حساب و گردش اعتبار مشتری

پس از حذف فیلدهای مالی، نتایج اعتبارسنجی الگوریتم‌ها نشان می‌دهد همچنان الگوریتم‌های شبکهٔ عصبی و درخت تصمیم، دقت بسیار مناسبی دارند و درخت تصمیم بهترین پاسخ را به یادگیری طبقه‌بندی و پیش‌بینی مشتری ویژه ارائه می‌کند. نتایج اعتبارسنجی روش بیزین پایه به‌طور مشخص کمتر بود و روش K نزدیک‌ترین همسایه نیز که جزء روش‌های طبقه‌بندی غیردقیق در پیش‌بینی و سریع در محاسبه به‌شمار می‌رود، دچار بیش‌برازش شده است که برای پیش‌بینی آینده مناسب نیست. نتیجه حاصل از درخت تصمیم در قالب شکل ۷ به نمایش گذاشته شده است.



شکل ۷. درخت تصمیم برای طبقه‌بندی و پیش‌بینی مشتریان ویژه بدون درنظرگیری ماندهٔ حساب و گردش بستانکار

بر اساس طبقه‌بندی به‌دست‌آمده از درخت تصمیم، به سهولت می‌توان گفت:

- جنسیت تأثیری در انتخاب مشتریان ویژه ندارد و تأثیر معناداری بر تصمیم‌گیری بانک نمی‌گذارد؛
- گردش بستانکار مناسب در حساب مشتری، احتمال تبدیل شدن وی به مشتری ویژه را افزایش می‌دهد؛

- متغیرهای بعدی که بسیار مهم هستند، سن و وضعیت تأهل مشتری است. اصولاً افراد میان‌سال متأهل، احتمال بیشتری دارد تا به مشتری‌های ویژه تبدیل شوند. مشتریانی که سن آنها حدود ۵۰ تا ۶۰ سال است، بازهٔ طلایی برای انتخاب مشتری ویژه خواهند بود.
- پس از آن، عدم وجود ادعای حقوقی علیه مشتری بیشترین احتمال را برای تبدیل‌شدن مشتری به مشتری ویژه ایجاد می‌کند.
- چک برگشتی تأثیر معناداری بر تبدیل‌شدن مشتری به مشتری ویژه ندارد؛ چراکه از طریق سازوکارهای نظام بانکی قابل پیگیری و رفع اثر است.
- نکتهٔ جالب دیگر، بی‌تأثیر بودن سطح تحصیلات و شغل مشتری بر تبدیل‌شدن به مشتری ویژه است که بنا به گفتهٔ خبرگان بانک منتخب، عمدتاً به‌دلیل بی‌دقتی زیاد این فیلدهای اطلاعاتی است؛ چراکه برخی از مشتریان ممکن است که شغل را عوض کنند یا از لحاظ تحصیلی به سطح بالاتری دست یابند، ولی به بانک اطلاع ندهند. با حذف این دو فیلد و طبقه‌بندی مجدد، مشخص شد که این دو فیلد تأثیر زیادی بر نتایج و یافته‌های طبقه‌بندی و پیش‌بینی مشتری ویژه ندارند.

ارزشمندی این یافته‌ها در آن است که تمام این موارد را در همان لحظهٔ ورود مشتری به بانک می‌توان تشخیص داد و سیاست‌های اجرایی و تعاملی بانک با مشتری را بر همان اساس تنظیم و هدایت کرد.

طبقه‌بندی و پیش‌بینی رتبهٔ مشتری ویژه

در مرحلهٔ دوم، با همان فیلدهای اطلاعاتی مرحلهٔ اول و بدون احتساب ماندهٔ حساب بانکی و گردش بستانکار مشتری، به طبقه‌بندی و یادگیری ویژگی‌های مشتریان بر اساس رتبهٔ مشتری ویژه اقدام شد. این سطوح در چهار طبقهٔ طلایی، نقره‌ای، پلاتینیوم و بدون سطح طبقه‌بندی شده‌اند. هدف از گام دوم داده‌کاوی این است که دریابیم آیا امکان پیش‌بینی رتبهٔ مشتری ویژه در آینده بر اساس داده‌های لحظهٔ ورود به بانک و تصمیم‌گیری بر اساس آن وجود دارد یا خیر. به همین دلیل، الگوریتم‌های منتخب طبقه‌بندی روی داده‌ها اعمال شدند و این فیلد به‌عنوان فیلد هدف برای یادگیری در نظر گرفته شد.

نتایج اعتبارسنجی یافته‌ها و درخت تصمیم به‌دست آمده در شکل‌های ۸ و ۹ مشاهده می‌شوند.

accuracy: 77.45%

	true NoGrade	true Platinium	true Silver	true Golden	class precision
pred. NoGrade	1189	60	0	20	93.70%
pred. Platinium	211	360	100	60	49.25%
pred. Silver	0	0	0	0	0.00%
pred. Golden	0	0	0	0	0.00%
class recall	84.93%	85.71%	0.00%	0.00%	

accuracy: 92.05%

	true NoGrade	true Platinium	true Silver	true Golden	class precision
pred. NoGrade	1391	80	0	0	94.56%
pred. Platinium	9	330	20	40	82.71%
pred. Silver	0	10	80	0	88.89%
pred. Golden	0	0	0	40	100.00%
class recall	99.36%	78.57%	80.00%	50.00%	

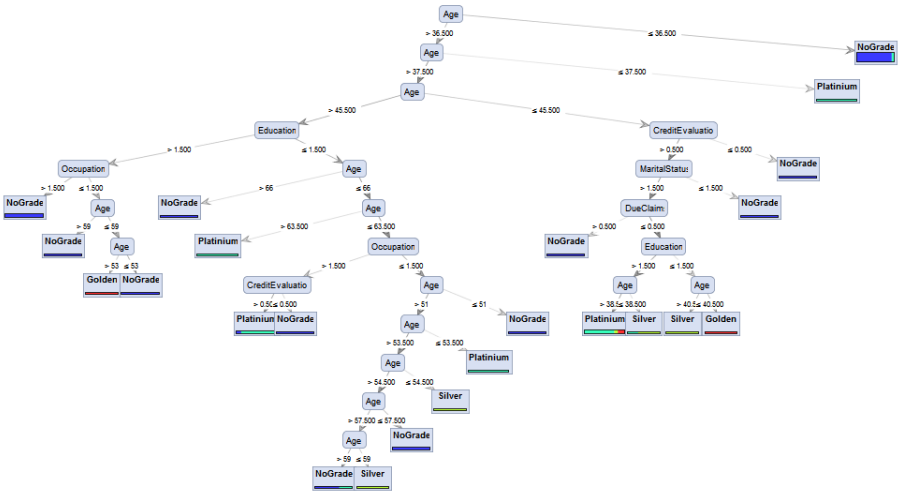
accuracy: 29.55%

	true NoGrade	true Platinium	true Silver	true Golden	class precision
pred. NoGrade	391	0	0	0	100.00%
pred. Platinium	530	100	0	0	15.87%
pred. Silver	267	140	60	40	11.83%
pred. Golden	212	180	40	40	8.47%
class recall	27.93%	23.81%	60.00%	50.00%	

accuracy: 92.50%

	true NoGrade	true Platinium	true Silver	true Golden	class precision
pred. NoGrade	1400	80	0	0	94.59%
pred. Platinium	0	330	20	40	84.62%
pred. Silver	0	10	80	0	88.89%
pred. Golden	0	0	0	40	100.00%
class recall	100.00%	78.57%	80.00%	50.00%	

شکل ۸. دقت الگوریتم‌های طبقه‌بندی برای پیش‌بینی سطح مشتری ویژه



شکل ۹. درخت تصمیم پیش‌بینی کنندهٔ رتبهٔ مشتری ویژه

بر مبنای دقت اعتبارسنجی‌ها و درخت تصمیم ایجاد شده، می‌توان دریافت کـه سـن، سـطح تحصیلات و ردهٔ شغلی، در تشخیص سطح مشتری ویژه مؤثرند. بر مبنای درخت تصمیم می‌توان مشخص کرد که هرچه سن و تا حدی تحصیلات بیشتر باشد، امکان افزایش سریع رتبهٔ مشتری ویژه وجود خواهد داشت. از سوی دیگر، گردش بستانکار مشـتری بلافاصـله پـس از تحصـیلات اثرگذار است. عدم وجود ادعای حقوقی در سال‌های گذشته نیز بر افزایش سریع سطح مشتری از پلاتینیوم به طلایی تأثیر می‌گذارد.

در ادامه، درختوارهٔ اگرـ آنگاه مربوط به شناسایی سطح مشتریان ویژه در بدو ورود بـه بانـک، ارائه شده است. همان‌گونه که مشاهده می‌شود، هرچند سن بسیار تأثیرگذار است با ورود جوانـان سرمایه‌گذار و کارآفرین به حوزهٔ کسب‌وکار، اثرگذاری سن بـر تبـدیل‌شـدن بـه مشـتریان ویـژه، کاهش می‌یابد.

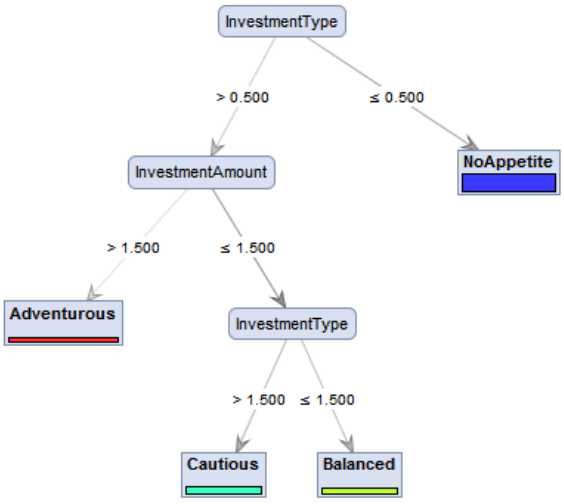
Tree

```
Age > 36.500
|   Age > 37.500
|   |   Age > 45.500
|   |   |   Education > 1.500
|   |   |   |   Occupation > 1.500: NoGrade {NoGrade=176, Platinium=0, Silver=0, Golden=0}
|   |   |   |   Occupation ≤ 1.500
|   |   |   |   |   Age > 59: NoGrade {NoGrade=46, Platinium=0, Silver=0, Golden=0}
|   |   |   |   |   Age ≤ 59
|   |   |   |   |   |   Age > 53: Golden {NoGrade=0, Platinium=0, Silver=0, Golden=20}
|   |   |   |   |   |   Age ≤ 53: NoGrade {NoGrade=36, Platinium=0, Silver=0, Golden=0}
|   |   |   Education ≤ 1.500
|   |   |   |   Age > 66: NoGrade {NoGrade=54, Platinium=0, Silver=0, Golden=0}
|   |   |   |   Age ≤ 66
|   |   |   |   |   Age > 63.500: Platinium {NoGrade=0, Platinium=40, Silver=0, Golden=0}
|   |   |   |   |   Age ≤ 63.500
|   |   |   |   |   |   Occupation > 1.500
|   |   |   |   |   |   |   CreditEvaluation > 0.500: Platinium {NoGrade=9, Platinium=60, Silver=0, Golden=0}
|   |   |   |   |   |   |   CreditEvaluation ≤ 0.500: NoGrade {NoGrade=18, Platinium=0, Silver=0, Golden=0}
|   |   |   |   |   |   Occupation ≤ 1.500
|   |   |   |   |   |   |   Age > 51
|   |   |   |   |   |   |   |   Age > 53.500
|   |   |   |   |   |   |   |   |   Age > 54.500
|   |   |   |   |   |   |   |   |   |   Age > 57.500
|   |   |   |   |   |   |   |   |   |   |   Age > 59: NoGrade {NoGrade=37, Platinium=20, Silver=0, Golden=0}
|   |   |   |   |   |   |   |   |   |   |   Age ≤ 59: Silver {NoGrade=0, Platinium=0, Silver=20, Golden=0}
|   |   |   |   |   |   |   |   |   |   Age ≤ 57.500: NoGrade {NoGrade=65, Platinium=0, Silver=0, Golden=0}
|   |   |   |   |   |   |   |   |   Age ≤ 54.500: Silver {NoGrade=0, Platinium=0, Silver=20, Golden=0}
|   |   |   |   |   |   |   |   Age ≤ 53.500: Platinium {NoGrade=0, Platinium=20, Silver=0, Golden=0}
|   |   |   |   |   |   |   Age ≤ 51: NoGrade {NoGrade=37, Platinium=0, Silver=0, Golden=0}
|   |   Age ≤ 45.500
|   |   |   CreditEvaluation > 0.500
|   |   |   |   MaritalStatus > 1.500
|   |   |   |   |   DueClaims > 0.500: NoGrade {NoGrade=10, Platinium=0, Silver=0, Golden=0}
|   |   |   |   |   DueClaims ≤ 0.500
|   |   |   |   |   |   Education > 1.500
|   |   |   |   |   |   |   Age > 38.500: Platinium {NoGrade=0, Platinium=170, Silver=20, Golden=40}
|   |   |   |   |   |   |   Age ≤ 38.500: Silver {NoGrade=0, Platinium=10, Silver=20, Golden=0}
|   |   |   |   |   |   Education ≤ 1.500
|   |   |   |   |   |   |   Age > 40.500: Silver {NoGrade=0, Platinium=0, Silver=20, Golden=0}
|   |   |   |   |   |   |   Age ≤ 40.500: Golden {NoGrade=0, Platinium=0, Silver=0, Golden=20}
|   |   |   |   MaritalStatus ≤ 1.500: NoGrade {NoGrade=27, Platinium=0, Silver=0, Golden=0}
|   |   |   CreditEvaluation ≤ 0.500: NoGrade {NoGrade=28, Platinium=0, Silver=0, Golden=0}
|   Age ≤ 37.500: Platinium {NoGrade=0, Platinium=40, Silver=0, Golden=0}
Age ≤ 36.500: NoGrade {NoGrade=857, Platinium=60, Silver=0, Golden=0}
```

طبقه‌بندی و پیش‌بینی میزان تمایل به ریسک مشتری ویژه

پس از امکان‌سنجی پیش‌بینی تبدیل مشتری به مشتری ویژه و همچنین پیش‌بینی رتبهٔ مشتری ویژه، این امکان نیز در داده‌ها فراهم آمده است که میزان تمایل به ریسک مشتریان سنجیده شده و توسط الگوریتم‌ها آموخته شود. به همین منظور متغیرهای مربوط به سرمایه‌گذاری شامل سطح و مدت زمان سرمایه‌گذاری، وجود تجربهٔ زیان مالی، میزان هدفمندی سرمایه‌گذاری و مبلغ سرمایه‌گذاری در بانک در نظر گرفته شدند. این فیلدها، متغیرهایی هستند که به‌طور معمول برای ارزیابی سطح ریسک‌پذیری مشتریان سرمایه‌گذار استفاده می‌شوند.

نتایج نشان داد میزان دقت الگوریتم‌ها در سنجش این متغیر نیز بسیار مناسب است و امکان تفکیک دقیق و صحیح میزان تمایل به ریسک‌پذیری مشتریان وجود دارد. درخت تصمیم حاصل از این ارزیابی در شکل ۱۰ مشاهده می‌شود.



شکل ۱۰. درخت تصمیم پیش‌بینی کنندهٔ میزان تمایل به ریسک مشتری ویژه

با توجه به هر یک از شاخه‌های اصلی درخت، می‌توان دریافت که سطح و مبلغ سرمایه‌گذاری، بیشترین تأثیر را در انتخاب مشتری ویژه در بانک دارد؛ به‌صورتی که مسائلی مانند تجربهٔ زیان مالی در گذشته، مدت زمان سرمایه‌گذاری و سطح هدف سرمایه‌گذاری، اثر چندانی بر میزان تمایل مشتری به ریسک‌پذیری نمی‌گذارند.

با توجه به اینکه متغیرهای ورودی توسط بانک امتیازدهی شده‌اند و هدف از پژوهش حاضر، تشخیص میزان تمایل به ریسک در بدو ورود به بانک است، بار دیگر میزان تمایل مشتریان به ریسک بر اساس متغیرهای اولیهٔ تحقیق بدون درنظرگیری متغیرهای سرمایه‌گذاری و با لحاظ متغیر تمایل به ریسک به‌عنوان متغیر خروجی، به‌کمک الگوریتم‌ها سنجیده خواهد شد. نتایج این اعتبارسنجی و درخت تصمیم نهایی در شکل‌های ۱۱ و ۱۲ مشاهده می‌شود.

accuracy: 97.40%

	true NoAppetite	true Cautious	true Balanced	true Adventurous	class precision
pred. NoAppetite	1400	14	3	0	98.80%
pred. Cautious	0	257	5	0	98.09%
pred. Balanced	0	29	251	0	89.64%
pred. Adventurous	0	0	1	40	97.56%
class recall	100.00%	85.67%	96.54%	100.00%	

accuracy: 99.95%

	true NoAppetite	true Cautious	true Balanced	true Adventurous	class precision
pred. NoAppetite	1399	0	0	0	100.00%
pred. Cautious	1	300	0	0	99.67%
pred. Balanced	0	0	260	0	100.00%
pred. Adventurous	0	0	0	40	100.00%
class recall	99.93%	100.00%	100.00%	100.00%	

accuracy: 65.10%

	true NoAppetite	true Cautious	true Balanced	true Adventurous	class precision
pred. NoAppetite	920	0	0	0	100.00%
pred. Cautious	111	102	0	0	47.89%
pred. Balanced	304	198	260	20	33.25%
pred. Adventurous	65	0	0	20	23.53%
class recall	65.71%	34.00%	100.00%	50.00%	

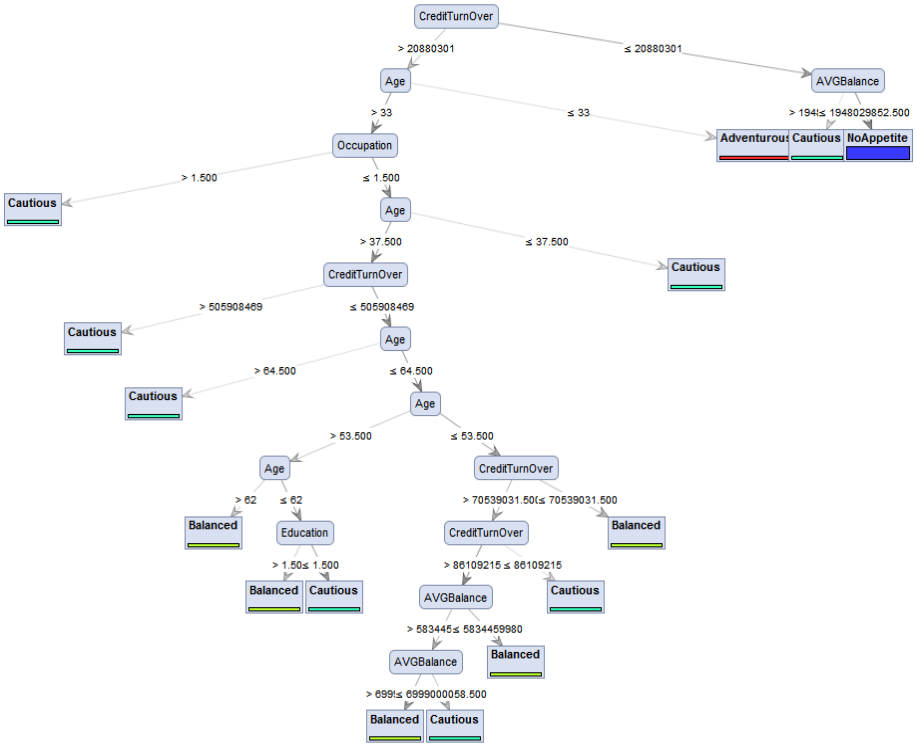
accuracy: 100.00%

	true NoAppetite	true Cautious	true Balanced	true Adventurous	class precision
pred. NoAppetite	1400	0	0	0	100.00%
pred. Cautious	0	300	0	0	100.00%
pred. Balanced	0	0	260	0	100.00%
pred. Adventurous	0	0	0	40	100.00%
class recall	100.00%	100.00%	100.00%	100.00%	

شکل ۱۱. دقت الگوریتم‌های طبقه‌بندی برای پیش‌بینی سطح میزان تمایل به ریسک مشتری ویژه

درخت تصمیم ارائه شده در شکل ۱۲، نشان‌دهندهٔ این مهم است که با دقت و اعتبار بسیار مناسبی، می‌توان به شناسایی میزان تمایل به ریسک مشتریان ویژه اقدام کرد. در حقیقت، گردش بستانکار مشتری، ماندهٔ حساب و سن مشتری، اثرگذارترین متغیرها هستند. با حذف

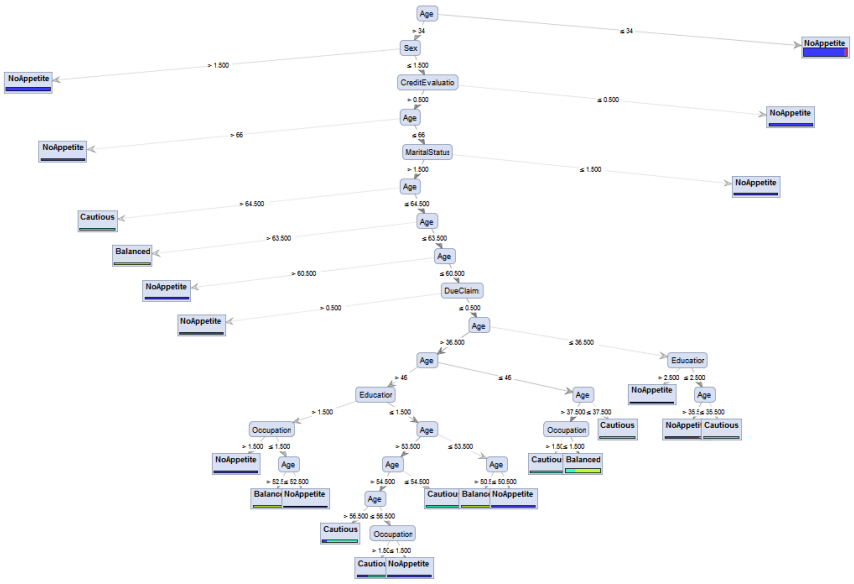
گردش بستانکار و ماندهٔ حساب، همچنان می‌توان دریافت که درخت تصمیم بسیاری دقیقی با اعتبار بیش از ۹۰ درصد خواهیم داشت که می‌تواند بر اساس شاخص‌های سن، جنسیت، ارزیابی اعتبار، وضعیت تأهل، دعاوی حقوقی، سطح تحصیلات و شغل، سطح تمایل مشتری به ریسک را مشخص کند (شکل ۱۳).



شکل ۱۲. درخت تصمیم پیش‌بینی‌کنندهٔ میزان تمایل به ریسک مشتری ویژه بر اساس ویژگی‌های اولیهٔ مشتری

accuracy: 92.60%

	true NoAppetite	true Cautious	true Balanced	true Adventurous	class precision
pred. NoAppetite	1382	10	0	40	96.51%
pred. Cautious	18	210	0	0	92.11%
pred. Balanced	0	80	260	0	76.47%
pred. Adventurous	0	0	0	0	0.00%
class recall	98.71%	70.00%	100.00%	0.00%	



شکل ۱۳. درخت تصمیم پیش‌بینی‌کنندهٔ میزان تمایل به ریسک مشتری ویژه بعد از حذف گردش اعتباری و ماندهٔ حساب مشتری

مجموعهٔ قواعد اگر ـ آنگاه مرتبط با درخت تصمیم شکل ۱۳ نیز در ادامه ارائه شده است.

```
Age > 34
|   Sex > 1.500: NoAppetite {NoAppetite=229, Cautious=0, Balanced=0, Adventurous=0}
|   Sex ≤ 1.500
|   |   CreditEvaluation > 0.500
|   |   |   Age > 66: NoAppetite {NoAppetite=46, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   Age ≤ 66
|   |   |   |   MaritalStatus > 1.500
|   |   |   |   |   Age > 64.500: Cautious {NoAppetite=0, Cautious=20, Balanced=0, Adventurous=0}
|   |   |   |   |   Age ≤ 64.500
|   |   |   |   |   |   Age > 63.500: Balanced {NoAppetite=0, Cautious=0, Balanced=20, Adventurous=0}
|   |   |   |   |   |   Age ≤ 63.500
|   |   |   |   |   |   |   Age > 60.500: NoAppetite {NoAppetite=55, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   Age ≤ 60.500
|   |   |   |   |   |   |   |   DueClaims > 0.500: NoAppetite {NoAppetite=10, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   DueClaims ≤ 0.500
|   |   |   |   |   |   |   |   |   Age > 36.500
|   |   |   |   |   |   |   |   |   |   Age > 46
|   |   |   |   |   |   |   |   |   |   |   Education > 1.500
|   |   |   |   |   |   |   |   |   |   |   |   Occupation > 1.500: NoAppetite {NoAppetite=46, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   Occupation ≤ 1.500
|   |   |   |   |   |   |   |   |   |   |   |   |   Age > 52.500: Balanced {NoAppetite=0, Cautious=0, Balanced=20, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   Age ≤ 52.500: NoAppetite {NoAppetite=18, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   Education ≤ 1.500
|   |   |   |   |   |   |   |   |   |   |   |   Age > 53.500
|   |   |   |   |   |   |   |   |   |   |   |   |   Age > 54.500
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Age > 56.500: Cautious {NoAppetite=9, Cautious=60, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   Age ≤ 56.500
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   Occupation > 1.500: Cautious {NoAppetite=9, Cautious=20, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   Occupation ≤ 1.500: NoAppetite {NoAppetite=38, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   Age ≤ 54.500: Cautious {NoAppetite=0, Cautious=40, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   Age ≤ 53.500
|   |   |   |   |   |   |   |   |   |   |   |   |   Age > 50.500: Balanced {NoAppetite=0, Cautious=0, Balanced=20, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   Age ≤ 50.500: NoAppetite {NoAppetite=19, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   Age ≤ 46
|   |   |   |   |   |   |   |   |   |   |   Age > 37.500
|   |   |   |   |   |   |   |   |   |   |   |   Occupation > 1.500: Cautious {NoAppetite=0, Cautious=20, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   |   Occupation ≤ 1.500: Balanced {NoAppetite=0, Cautious=80, Balanced=200, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   Age ≤ 37.500: Cautious {NoAppetite=0, Cautious=40, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   Age ≤ 36.500
|   |   |   |   |   |   |   |   |   |   Education > 2.500: NoAppetite {NoAppetite=19, Cautious=0, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   Education ≤ 2.500
|   |   |   |   |   |   |   |   |   |   |   Age > 35.500: NoAppetite {NoAppetite=36, Cautious=10, Balanced=0, Adventurous=0}
|   |   |   |   |   |   |   |   |   |   |   Age ≤ 35.500: Cautious {NoAppetite=0, Cautious=10, Balanced=0, Adventurous=0}
|   |   |   |   MaritalStatus ≤ 1.500: NoAppetite {NoAppetite=46, Cautious=0, Balanced=0, Adventurous=0}
|   |   CreditEvaluation ≤ 0.500: NoAppetite {NoAppetite=74, Cautious=0, Balanced=0, Adventurous=0}
Age ≤ 34: NoAppetite {NoAppetite=746, Cautious=0, Balanced=0, Adventurous=40}
```

در صورتی که بانک بتواند بر اساس شرایط اولیهٔ مشتری و بدون سرمایه‌گذاری، سطح مشتری ویژه و میزان تمایل وی به ریسک را معین کند، سیاست‌گذاری و تصمیم‌گیری تعاملی بانک به سمت مشتریان ویژه با دقت بسیار زیادی امکان‌پذیر می‌شود که در طولانی‌مدت، به بهبود سطح عملکرد مالی و اعتباری بانک در کشور می‌انجامد. به همین منظور، پیشنهاد می‌شود سیستمی طراحی شود که بر مبنای شاخص‌ها و قواعد اگر- آنگاه ارائه شده در بخش قبل و همچنین مجموعه قواعد اگر- آنگاه ارائه شده برای درخت تصمیم شکل ۱۳، به‌طور مستقیم قادر به شناسایی مشتری ویژه، رتبهٔ امتیاز مشتری و میزان تمایل وی به ریسک باشد و بر مبنای آن، از همان لحظهٔ اول مواجهه با مشتری، سیاست‌گذاری‌های تعاملی لازم را انجام دهد و مشتری را با سرعت بیشتری تبدیل به مشتری وفادار کند.

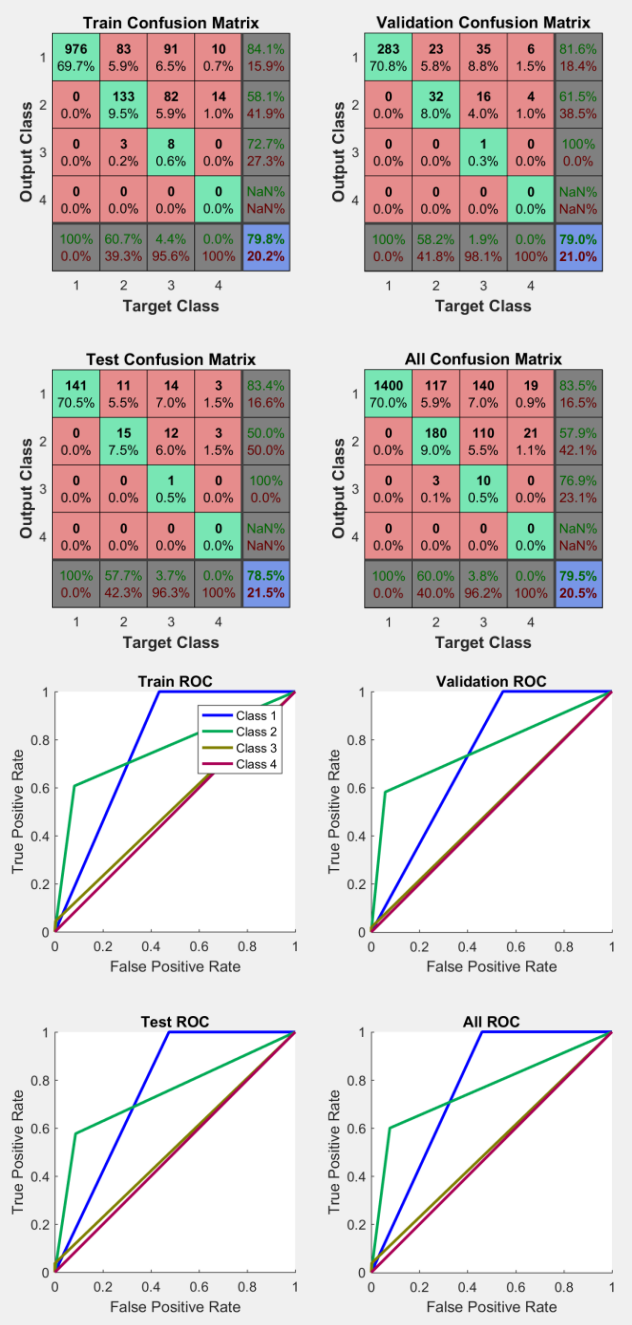
بررسی تمایل به ریسک توسط کمی‌سازی بردارهای یادگیرنده[1]

در نهایت به‌منظور اطمینان از کیفیت طبقه‌بندی و پیش‌بینی انجام شده، مقایسه‌ای نیز بین نتایج به‌دست‌آمده از چهار الگوریتم بخش قبل با نتایج حاصل از الگوریتم کمی‌سازی بردارهای یادگیرنده صورت گرفت. این الگوریتم یکی از روش‌های نوین و دقیق برای طبقه‌بندی داده‌های کمی است که انواع کلاس‌های خروجی را در قالب بردارهای متشکل از صفر و یک طبقه‌بندی می‌کند؛ سپس یادگیری را توسط شبکه‌های عصبی انجام می‌دهد. برای داده‌های این پژوهش، از روش بردار یادگیرنده، نسخهٔ ۲/۱ که نسخهٔ جدید این روش است، استفاده می‌شود. نتایج به‌دست آمده نشان داد با این که تعداد داده‌ها برای این الگوریتم به‌طور نسبی کم بود، قابلیت پیش‌بینی مناسبی از الگوریتم به‌دست آمد و دقتی نزدیک به ۸۰ درصد در پیش‌بینی میزان تمایل به ریسک مشتری در بدو ورود به سازمان حاصل شد (شکل ۱۴).

این‌گونه می‌توان نتیجه گرفت که مجموعهٔ الگوریتم‌های شبکهٔ عصبی (پرسپترون چندلایه و بردارهای یادگیرنده)، هرچند برای بردارهای گسسته نیز بهینه‌سازی شده‌اند، در مقایسه با الگوریتم‌های درخت تصمیم، بیزین پایه و نزدیک‌ترین همسایه، دقت نسبتاً کمتری دارند و نیاز به بهینه‌سازی بیشتری در آنها احساس می‌شود. با وجود این، نتایج به‌دست آمده حاکی از دقت بسیار مناسب مجموعه الگوریتم‌ها در پیش‌بینی رتبه و تمایل به ریسک مشتریان است و تمام الگوریتم‌ها را می‌توان به‌صورت رقابتی برای دستیابی به نتیجهٔ مطلوب و پیش‌بینی دقیق رویکرد مشتریان نسبت به بانک به خدمت گرفت.

---

1. Learning Vector Quantization (LVQ)



شکل ۱۴. اعتبارسنجی الگوریتم کمی‌سازی بردار یادگیرنده

نتیجه‌گیری

با توجه به نیاز بانک‌های کشور به شناسایی و تفکیک مشتریان ویژه و ممتاز نسبت به مشتریان معمولی و پرتکرار، در این پژوهش تلاش شد با استفاده از مجموعه‌ای از داده‌های گردآوری شده از یکی از بانک‌های بزرگ کشور، نسبت به طراحی سازوکاری برای شناسایی هوشمند و خودکار این مشتریان در بدو ورود به بانک و تفکیک نوع و کیفیت خدمت قابل ارائه به آنها در مقایسه با سایر مشتریان اهتمام شود. با شناسایی این مشتریان و تشخیص رتبهٔ احتمالی آنها در آینده و همچنین میزان تمایل آنها به ریسک، می‌توان خدمات متناسب با کیفیت و کمیت سرمایه‌گذاری مشتریان ویژه در بانک ایجاد و ارائه کرد. در ادامه، به برخی از خدماتی که در حفظ و ایجاد وفاداری این مشتریان تأثیرگذارند، اشاره می‌شود: صدور کارت‌های خاص برای مشتریان ویژه؛ شناسایی نیازهای کوتاه‌مدت و بلندمدت مشتریان ویژه؛ طراحی بسته‌های خدمات کامل متناسب با مشتریان ویژه؛ تحلیل مدیریت ریسک مشتریان ویژه و ارائه به آنها برای کاهش ریسک سرمایه‌گذاری؛ مستندسازی و به‌روزرسانی پروفایل مشتری از طریق جمع‌آوری اطلاعاتی همچون وضعیت مالی، میزان تحمل ریسک، اهداف و افق زمانی سرمایه‌گذاری؛ محدودیت‌هایی مانند ترجیحات سرمایه‌گذاری و نیازهای نقدینگی به‌منظور ارائهٔ خدمات دقیق و کارآمد به این مشتریان؛ ملاقات با مشتریان در بازه‌های زمانی متفاوت؛ ارسال هدیه برای مناسبت‌های مختلف به مشتریان ویژه؛ تسهیل فرایند پاسخگویی به نیازهای این مشتریان؛ ارائهٔ خدمات و تسهیلات و طرح‌های خاص به مشتریان ویژه در اسرع وقت و بدون قرارگیری در نوبت‌های طولانی؛ افزایش سود بانکی سپرده‌ها و حذف کارمزد بانکی برای مشتریان ویژه؛ استخدام مشاوران خبرهٔ مالی و ارائهٔ مشاوره به مشتریان ویژه در امور متنوع مالی همانند مشاوره‌های خرید و فروش سهام، ارز و طلا، مدیریت املاک مشتریان؛ ارائه خدمات و تسهیلات در محل کار مشتریان ویژه؛ افزایش سرمایهٔ مشتریان ویژه از طریق برنامه‌های مستمر و ترغیب‌کننده؛ صدور کارت اعتباری ارزی برای مشتریان اختصاصی و ممتاز؛ خدمات ارزی و صرافی؛ ارائهٔ خدمات در زمینهٔ پزشکی و بیمهٔ تکمیلی درمان؛ ارائهٔ خدمات در زمینهٔ روزو هتل و بیمهٔ مسافرتی؛ ارائهٔ خدمات در زمینهٔ سفارش اینترنتی، ارسال هدایا و سبد گل؛ ارائهٔ خدمات در زمینهٔ برگزاری جشن‌ها و مراسم؛ ارائهٔ خدمات در زمینهٔ مشاورهٔ حقوقی؛ خلاقیت و نوآوری در ارائهٔ خدمات انجام درخواست‌های روزانهٔ مشتریان همانند رزور بلیط هواپیما، ارتقا کیفیت خدمات ارائه شده توسط شرکت‌های زیرمجموعهٔ بانکداری اختصاصی و ارائهٔ آموزش‌های ویژه به مشتریان همانند آموزش فرایندها و قوانین صادرات و واردات کالا و خدمات و سایر خدماتی که می‌توانند به بهبود کیفیت همکاری مشتری ویژه با بانک منجر شوند و بر وفاداری وی نسبت به بانک بیفزایند.

همچنین در زمینهٔ تحقیقات آتی پیشنهاد می‌شود که از الگوریتم‌های تکاملی به‌منظور بهبود نتایج الگوریتم‌های فعلی استفاده شود. از سوی دیگر، پیشنهاد می‌شود که با بهره‌مندی از اطلاعات مشتریان ویژهٔ بانک‌های دیگر، مقایسه و تحلیل تطبیقی نسبت به رویکرد بانک‌های مختلف در برابر مشتریان ویژه ارائه شود و مشتریان ارزی نیز به‌صورت مجزا مدنظر قرار گیرند. همچنین می‌توان عملکرد بانک‌ها را در یک بازهٔ زمانی نسبت به هزینهٔ ارائهٔ خدمات به مشتریان ویژه و درآمد کسب شده از طریق آنها، تحلیل کرد و با استفاده از الگوریتم‌های تکاملی مختلف، نسبت به شناسایی نقطهٔ بهینهٔ ارائهٔ خدمات به مشتریان ویژه اقدام نمود.

فهرست منابع

اسماعیل‌نیا، م. (۱۳۹۱). ردپای بانکداری اختصاصی و شرکتی در برنامه‌های سال ۹۱ بانک‌های خصوصی، *نشریهٔ تبلیغات بازرگانی*، (۵۱ و ۵۲)، ۲۹-۲۶.

بازرگان، ع.، سرمد، ز. و حجازی، ا. (۱۳۸۶). *روش‌های تحقیق در علوم رفتاری*. چاپ سیزدهم، تهران: انتشارات آگاه.

توتونچیان، ا. (۱۳۷۵). *اقتصاد پول و بانکداری*. تهران: مؤسسهٔ تحقیقات پولی و بانکی.

دلاور، م. (۱۳۹۲). *روش تحقیق در روان‌شناسی و علوم تربیتی*، انتشارات دانشگاه پیام‌نور.

رضایی‌نور، ج.، شیخ بهایی، ج. (۱۳۹۶). کاربردهای داده‌کاوی متنی در حوزهٔ مدیریت دانش زنجیرهٔ خدمات دولت الکترونیکی. *نشریهٔ مدیریت فناوری اطلاعات*، ۹ (۱)، ۶۶-۳۹.

رئیسی وانانی، ا.، گنجعلی خان حاکمی، ف. (۱۳۹۴). طراحی سیستم استنتاج فازی ـ عصبی انطباقی برای ارزیابی استقرار سیستم هوشمندی کسب‌وکار در صنعت تولید نرم‌افزار. *نشریهٔ مدیریت فناوری اطلاعات*، ۹ (۱)، ۸۵-۱۰۴.

شفیعی رود پشتی، م.، حکاکی، س.م.، جلالی، م.، نوری، ا. (۱۳۹۳). تحلیل بر وضعیت بانکداری اختصاصی. *فصلنامهٔ پژوهش و سیاست‌های اقتصادی*، ۲۲(۷۰)، ۱۳۸- ۱۱۹.

شهرابی ج.، شجاعی، ع.ذ. (۱۳۹۰). *داده‌کاوی پیشرفته: مفاهیم و الگوریتم‌ها*. تهران: جهاد دانشگاهی واحد صنعتی امیرکبیر.

طاهری، ن. (۱۳۹۱). کلیاتی از بانکداری اختصاصی (Private Banking). *نشریهٔ بانک و اقتصاد*، ۱۱۹، ۴۰-۳۶.

غلامپور فرد، م.م.، مختاری، م.، رضوی خسروانی، آ. (۱۳۹۱). بانکداری به میل مشتری، نگاهی به بانکداری اختصاصی در جهان. *فصلنامهٔ تازه‌های اقتصاد*، ۱۳۸، ۱۵۱-۱۴۹.

منصف، ع.، منصوری، ن. (۱۳۸۹). بررسی عوامل مؤثر بر حجم سپرده‌های بانکی (با تأکید بر نرخ سود اوراق مشارکت: ۱۳۸۷-۱۳۶۷)، *مجلهٔ دانش و توسعه*، ۱۷ (۳۴)، ۹۰-۶۹.

ناظمی یگانه، ف. (۱۳۸۹). بانکداری اختصاصی، حلقهٔ گمشدهٔ نظام بانکی کشور. *فصلنامهٔ تخصصی بانک صادرات*، ۱۱(۵۴)، ۵۰-۴۷.

Ansari, A. & Riasi, A. (2016). Modelling And evaluating customer loyalty using neural networks: Evidence from startup insurance companies. *Future Business Journal,* 2(1), 15-30.

Bazargan, A., Sarmad, Z. & Hejazi, A. (2007). *Research Methods in Behvioral Sciences* (13th Edition). Tehran: Agah Publications. (*in Persian*)

Carrasco, I. G., Palacios, R. C., Cuadrado, J. L., Crespo, A. G. & B. R. Mezcua (2012). PB-ADVISOR: A Private Banking Multi-Investment Portfolio Advisor. *Information Sciences*, 206, 63–82.

Delavar, M. (2013). *Research Method in Psychology and Educational Sciences*. Payam-e-Noor University Publication. (*in Persian*)

Dichev, C. & Dicheva, D. (2017). Towards Data Science Literacy, Procedia Computer Science 108C. *International Conference on Computational Science,* 2151–2160.

Dömötör, B. (2011). A kockázat megjelenése a származtatott pénzügyi termékekben (The risk appearing in financial derivatives). *Credit Institution Review*, 10(4), 360– 369.

Dömötör, B. (2013). Modelling optimal Hedge Ratio in the Presence of Funding Risk. *27th European conference on Modelling and simulation*, 27–30.

Esmailnia, M. (2012). The trace of private and corporate banking in the plans of private banks in 1391. *Business Marketing Journal*, (51- 52), 26-29. (*in Persian*)

Gallan, A.S., Jarvis, C.B., Brown, S.W. & Bitner, M.J. (2013). Customer positivity and participation in services: An empirical test in a healthcare context. *Journal of the Academy of Marketing Science*, 41(3), 338–356.

Gholampour Fard, M.M., Mokhtari, M. & Razavi Khosravani, A. (2012). Customer-Oriented Banking: A look on Private Banking in the World. *Economic News Journal*, 138, 149-151. (*in Persian*)

Han, J., Kamber, M. & Pei, J. (2011). *Data Mining: Concepts and Techniques (Third Edition). SanFrancisco.* Morgan Kaufmann Publishers, CA, USA.

Han, J., Kamber, M. (2006). *Data Mining concepts and techniques*. Second Edition, Morgan Kaufmann.

Kirakosyan, K. & Dănăiață, D. (2014). Communication management in electronic banking. Better communication for better relationship. *Procedia-Social and Behavioral Sciences,* 124(2014), 361–370.

Larose, D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining*. 2nd Edition. John Wiley-Interscience.

Li, X. & Petrick, J.F. (2008). Reexamining the dimensionality of brand loyalty: A case of the cruise industry. *Journal of Travel Tourism Marketing*, 25(1), 68–85.

Liao, S.H., Chu, P.H. & Hsiao, P.Y. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications,* 39(12), 11303-11311.

Liu, C. M. (2001), An Assessment of Banking Operation Strategies of Private Banking Institutions in the Philippines. *Asia Pacific Journal of Marketing & Logistics*, 13(1), 57-71.

Markov, Z. & Rusell, I. (2009) .An Introduction to the WEKA data mining system. *Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education*, 367-368.

Monsef, A. & Mansouri, N. (2010). Evaluation of Influential Factors on the Volume of Bank Deposits (With Respect to the Profit Rate of Bonds: 1367-1387). *Knowledge and Development Journal*, 17(34), 69-90. (*in Persian*)

Nazemi Yeganeh, F. (2010). Private Banking: Missing Ring of Banking System in the Country. *Journal of Saderat Bank*, 11(54), 47-50. (*in Persian*)

Raeesi Vanani, I. & Ganjalikhan Hakemi, F. (2015). Design of Adaptive Neuro-Fuzzy Inference System for Evaluating the Implementation of Business Intelligence System in Software Development Industry. *Information Technology Management Journal*, 9(1), 85-104. (*in Persian*)

Reydet, S. & Carsana, L. (2017). The effect of digital design in retail banking on customers' commitment and loyalty: The mediating role of positive affect. *Journal of Retailing and Consumer Services*, 37, 132-138.

Rezaee Noor, J. & Sheikh Bahayi, J. (2017). Applications of Textual Data Mining in Knowledge Management Domain of e-Government Service Chain. *Information Technology Management Journal*, 7(1), 39-66. (*in Persian*)

Shafiee Roodposhti, M., Hakaki, S.M., Jalali, M. & Noori, A. (2014). Analysis of Private Banking Conditions. *Journal of Economic Research and Policies,* 22(70), 119-138. (*in Persian*)

Shahrabi, J. & Shojaee, A.Z. (2011). *Advanced Data Mining, Concepts and Algorithms*. Tehran: University of Amirkabir, Jahad Daneshgahi Publication. (*in Persian*)

Straker, K., Wrigley, C. & Rosemann, M. (2015). The role of design in the future of digital channels: conceptual insights and future research directions. *Journal of Retailing and Consumer Services*, 26, 133–140.

Straker, K., Wrigley, C. (2016). Translating emotional insights into digital channel designs: opportunities to enhance the airport experience. *Journal of Hospitality and Tourism Technology,* 7 (2), 135–157.

Taheri, N. (2012). Summaries of Private Banking. *Economy and Bank Journal*, 119, 36-40. (*in Persian*)

Tootoonchian, A. (1996). *Money and Banking Economy*. Tehran: Banking and Money Research Institute. (*in Persian*)

Witten, I.H. & Frank, E. (2011). *Data Mining: Practical machine Learning tools and techniques*. (3 ed.). Elsevier.

Wrigley, C. (2013). Design dialogue: the visceral hedonic rhetoric framework. *Design Issues,* 29 (2), 82–95.