



## Designing and Explaining a Model for Profitable Automated Teller Machines (ATMs) Predicting and Locating of Tejarat-Bank in Hamadan Based on Security and Theft Prevention Approach using Data Mining Algorithms

### Orkideh Hamedei

\*Assistant Professor and Faculty Member, Department of Management, Economics and Accounting, Payame Noor University, Tehran, Iran.

orkideh.hamedei@gmail.com.

(Corresponding Author)

### Mohammad Amin Torabi

Department of Business Management, Faculty of Management, Economics and Accounting, Payame Noor University, Tehran, Iran.

### Mohammad Rafiei Nia

Master of Computer Engineering, Artificial Intelligence, Faculty of Computer Engineering, Islamic Azad University, Hamadan, Iran.

### Navid Esfandiari

Department of Management, Economics and Accounting, Payame Noor University, Tehran, Iran.

Received: 2020/12/17

Accepted: 2021/02/06

DOI:

10.22034/HPSJ.2021.95883

### ABSTRACT

Nowadays, use of automated teller machines (ATMs) have revolutionized the whole banking processes. There is no need for anyone to linger in long queues at the bank to receive money due to ATMs existence. ATM Installation at one site isn't risk-free and could be dangerous for banks in terms of security; therefore, bank security can be increased and theft can be prevented by the right and appropriate security locating of the ATM in a city. In this study, we investigate the prediction of the ATMs number required for a bank and their right locating using various data mining methods such as k-nearest neighbor's algorithm (k-NN), Rough Sets, classification tree algorithm, regression and neural network. The results of this study illustrate that the use of data mining tools can help officials in predicting and locating the required number of ATMs in Hamadan and prevent theft, increase security, and improve police security activities. In addition to considering the profitability of ATMs, the security coefficient of locating each branch is also considered in the data mining models considered in this study. According to the results, it is very important to pay attention to the following indicators in decision making, predicting number and ATMs locating: contract duration, total commission, total transactions and number of ATMs in each district, theft rate in the district, distance to the nearest police station, number of CCTV cameras in the district, population density and number of guards. Also, these indicators will be the basis for predicting the number of profitable ATMs. The support vector machines (SVMs) algorithm with root mean square error (RMSE) of 8% has better estimation capability than other algorithms.

**Keywords:** Data Mining, ATM Security, Theft Prevention.

► **Citation (Vancouver):** Hamedei O, Amin Torabi M, Rafiei Nia M, Esfandiari N. Designing and Explaining a Model for Profitable Automated Teller Machines (ATMs) Predicting and Locating of Tejarat-Bank in Hamadan Based on Security and Theft Prevention Approach using Data Mining Algorithms. *Quarterly J Hamedan Police Sci.* Winter 2021; 7(4):1-18.

► **Citation (APA):** Hamedei, O., Amin Torabi, M., Rafiei Nia, M., Esfandiari, N. (Winter 2021). Designing and Explaining a Model for Profitable Automated Teller Machines (ATMs) Predicting and Locating of Tejarat-Bank in Hamadan Based on Security and Theft Prevention Approach using Data Mining Algorithms. *Quarterly Journal of Hamedan Police Science*, 7(4), 1-18.

## طراحی و تبیین مدلی جهت پیش‌بینی و مکان‌یابی خودپردازهای سودده بانک تجارت شهر همدان با رویکرد امنیت و پیشگیری از سرقت با استفاده از الگوریتم‌های داده‌کاوی

### چکیده

امروزه، دستگاه‌های خودپرداز عملیات بانکی را متحول کرده‌اند. با وجود این دستگاه‌ها، دیگر لازم نیست کسی برای دریافت پول در صف‌های طولانی بانک معطل شود. نصب این دستگاه در یک مکان برای بانک‌ها می‌تواند از نظر امنیتی خطرناک باشد. از این رو، مکان‌یابی صحیح امنیتی و مناسب خودپرداز در یک شهر می‌تواند باعث افزایش امنیت بانک و جلوگیری از سرقت شود. در این مقاله با استفاده از روش‌های مختلف داده‌کاوی مانند الگوریتم نزدیک‌ترین همسایه، رافت، الگوریتم درخت طبقه‌بندی، رگرسیون و شبکه عصبی به پیش‌بینی تعداد دستگاه‌های خودپرداز موردنیاز یک بانک و جانمایی صحیح آن می‌پردازیم. با توجه به نتایج به دست آمده از تحقیق، دریافتیم که استفاده از ابزارهای داده‌کاوی می‌تواند به مسئولین در پیش‌بینی و جانمایی تعداد خودپرداز موردنیاز در شهر همدان کمک کند و باعث جلوگیری از سرقت، افزایش ضریب امنیتی و بهبود فعالیت‌های امنیتی پلیس گردد. در این پژوهش، علاوه بر در نظر گرفتن سوددهی خودپردازها، ضریب امنیتی مکان‌یابی هر شعبه را نیز در الگوریتم‌های داده‌کاوی مدنظر قرار دادیم. نتایج نشان داد که توجه به شاخص‌های مدت قرارداد، کل کارمزد، کل تراکنش‌ها و تعداد خودپردازها در هر منطقه، میزان سرقت در منطقه، مسافت نزدیک‌ترین کلانتری، تعداد دوربین‌های مداربسته منطقه، تراکم جمعیت و تعداد نهبان‌ها، در تصمیم‌گیری، پیش‌بینی تعداد و جانمایی خودپردازها بسیار حائز اهمیت است. این شاخص‌ها مبنایی برای پیش‌بینی تعداد خودپردازهای سودده نیز خواهند بود. همچنین، الگوریتم ماشین بردار پشتیبان با جذر میانگین مربعات خطای ۸ درصد، قابلیت برآورد بهتری را نسبت به سایر الگوریتم‌ها دارد.

**کلیدواژه‌ها:** داده‌کاوی، امنیت خودپردازها، جلوگیری از سرقت.

### ارکیده حامدی

\*استادیار، گروه مدیریت بازرگانی، دانشکده مدیریت، اقتصاد و حسابداری، دانشگاه پیام نور، تهران، ایران.

orkideh.hamed@gmail.com

(نویسنده مسئول)

### محمدامین ترابی

گروه مدیریت بازرگانی، دانشکده مدیریت، اقتصاد و حسابداری، دانشگاه پیام نور، تهران، ایران.

### محمد رفیعی‌نیا

کارشناسی ارشد مهندسی کامپیوتر، گرایش هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، همدان، ایران.

### نوید اسفندیاری

گروه مدیریت بازرگانی، دانشکده مدیریت، اقتصاد و حسابداری، دانشگاه پیام نور، تهران، ایران.

### نوع مقاله: پژوهشی

صص: ۱-۱۸

تاریخ دریافت: ۱۳۹۹/۰۹/۲۷

تاریخ پذیرش: ۱۳۹۹/۱۱/۱۸

شناسه دیجیتال (DOI):

10.22034/HPSJ.2021.95883

◀ **استناد (ونکوور):** حامدی، ا.، ترابی، م.، رفیعی‌نیا، م.، اسفندیاری، ن. طراحی و تبیین مدلی جهت پیش‌بینی و مکان‌یابی خودپردازهای سودده بانک تجارت شهر همدان با رویکرد امنیت و پیشگیری از سرقت با استفاده از الگوریتم‌های داده‌کاوی. فصلنامه علمی دانش انتظامی همدان. زمستان ۱۳۹۹؛ ۷(۴): ۱-۱۸.

◀ **استناد (APA):** حامدی، ا.، ترابی، م.، رفیعی‌نیا، م.، اسفندیاری، ن. (زمستان ۱۳۹۹). طراحی و تبیین مدلی جهت پیش‌بینی و مکان‌یابی خودپردازهای سودده بانک تجارت شهر همدان با رویکرد امنیت و پیشگیری از سرقت با استفاده از الگوریتم‌های داده‌کاوی. فصلنامه علمی دانش انتظامی همدان، ۷(۴)، ۱-۱۸.

در عصر حاضر، بانکداری الکترونیک در تعاملات پولی و مالی را می‌توان از مهم‌ترین و ضروری‌ترین ساختارهای صنعت نوین بانکداری به حساب آورد که شامل مفاهیم نوآورانه‌ای تحت عنوان پول الکترونیک<sup>۱</sup>، ماشین‌های خودپرداز<sup>۲</sup>، بانکداری خانگی<sup>۳</sup>، تلفن‌بانک<sup>۴</sup>، موبایل‌بانک<sup>۵</sup>، بانکداری اینترنتی<sup>۶</sup> و بانکداری مجازی<sup>۷</sup> است (صادقی، ۱۳۹۱). خودپرداز<sup>۸</sup> که گاه عابربانک هم نامیده شده دستگاهی الکترونیکی است که به مشتریان بانک این امکان را می‌دهد تا در هر زمان دلخواه به وسیله قرار دادن کارت الکترونیکی خاصی در دستگاه و واردکردن یک گذرواژه از حساب خود پول دریافت کرده یا موجودی حساب بانکی خود را دریافت کنند، بدون این‌که نیاز به شمارشگر انسانی باشد. بسیاری از خودپردازها نیز به امکان واریز پول یا چک، جابجایی پول میان حساب‌های بانکی و حتی خرید تمبر را برای مشتریان فراهم می‌کنند. نخستین دستگاه خودپرداز خاورمیانه، در ایران به وسیله بانک بیمه بازرگانان و در تهران قرار گرفت (کیهان، ۱۳۵۱). طبق آمارهای موجود، خودپرداز محبوب‌ترین روش پرداخت است. با توجه به پیشرفت سریع در تکنولوژی مدرن، استفاده از کارت‌های خودپرداز افزایش یافته و این باعث شده است بانک‌ها تعداد خودپردازهای خود را در سطح مناطق مختلف بیشتر کنند (رحمان و سها، ۲۰۱۹).

امروزه، اغلب دستگاه‌های خودپرداز ایرانی به شبکه شتاب متصل هستند. شبکه شتاب (شبکه تبادل اطلاعات بین‌بانکی) امکان برداشت و مشاهده موجودی را برای دارندگان کارت‌بانک‌های عضو این شبکه از همه خودپردازها فراهم می‌آورد. پس از اتصال بانک‌ها توسط شبکه شتاب در سال ۱۳۸۱ و ارائه خدمات بانک‌ها به

مشتریان یکدیگر، قوانینی وضع شد. شبکه شتاب باهدف یکپارچه کردن و به هم پیوستن سامانه‌های کارت تمامی بانک‌های شهر آغاز به کار کرد و در حال حاضر به‌عنوان درگاهی یکپارچه برای ارتباط مالی بانک‌های عضو این شبکه، در کل کشور فعال است (بانک مرکزی، ۱۳۸۱). مشابه هر کسب‌وکار و بنگاه تجاری، اساسی‌ترین هدف راهبردی بانک‌ها نیز، دستیابی به سود است. سودآوری را می‌توان جبهه‌ای مستحکم در مقابل بحران‌های اقتصادی بازار و تأثیرات منفی آن دانست که می‌تواند ثبات اقتصادی نظام مالی را تقویت کند (نوری‌بروجردی، جلیلی و مردانی، ۱۳۸۹). هر بانک، بسته به این‌که چه تعداد خودپرداز نصب شده و چه تعداد کارت صادر شده دارد، درنهایت مبالغی به‌عنوان کارمزد پرداخت و دریافت می‌نماید و پس از کسر این مبلغ، درنهایت بدهکار یا بستانکار شبکه بانکی می‌شود. بانک‌هایی که تعداد بیشتر و گستردگی مناسب‌تر شبکه خودپرداز دارند، ممکن است بستانکار شده و بانک‌هایی که به نسبت کارت‌های بیشتری صادر کرده‌اند و دستگاه‌های کم‌تری دارند، ممکن است بدهکار شوند. این مبالغ بانک‌ها را بر آن داشت که نسبت به نصب دستگاه‌های بیشتر اقدام نمایند (بانک مرکزی، ۱۳۸۱). دانستن تعداد خودپرداز موردنیاز در شعب بانک باعث کسب سود و نصب تعداد بیش‌ازاندازه به دلیل هزینه‌های خودپرداز مانند اجاره دستگاه، برق مصرفی و غیره باعث زیان‌دهی خواهد بود. برای آگاهی از سوددهی یا زیان‌دهی هر خودپرداز از کارمزد حاصل شده در طول ماه استفاده می‌شود. به دلیل حجم بالای اطلاعات در طول ماه‌های متوالی، تخمین تعداد خودپرداز موردنیاز در هر شهر و یا هر منطقه امری مشکل است.

6. Internet Banking  
7. Virtual Banking  
8. ATM  
9. Rahman & Saha

1. Digital Money  
2. Automated Teller Machine(ATM)  
3. Home Banking  
4. Telephone Banking  
5. Mobile Banking

امنیت خودپردازها از جمله موارد بااهمیتی است که بانکها بایستی در کنار بررسی سوددهی خودپردازها، در نظر بگیرند و بر مبنای عوامل بازدارنده اقدام به جانمایی صحیح امنیتی آن بکنند (اصغری و توانگر، ۱۳۹۱).

اگرچه این دستگاهها دارای خصوصیات و ویژگیهای مهمی برای ضد سرقت بودن میباشند، اما بایستی بانکها با همکاری نیروی انتظامی شهر مربوطهشان، شرایط لازم برای جانمایی بهینه را پیدا کنند و شعبه‌ای را در نظر بگیرند که در کنار سوددهی بالا، بیشترین ضریب امنیتی جهت جلوگیری از سرقت را به آن‌ها بدهد.

داده‌کاوی<sup>۱</sup> فرآیندی است خودکار برای استخراج الگوهایی که دانش را بازنمایی می‌کنند؛ این دانش به صورت ضمنی در پایگاه داده‌های عظیم، انباره داده<sup>۲</sup> و دیگر مخازن بزرگ اطلاعات، ذخیره شده است. به‌طورکلی، داده‌کاوی (که گاهی اکتشاف داده یا دانش هم نامیده می‌شود)، برای فرایند تحلیل داده‌ها از منظرهای متفاوت و خلاصه کردن آن‌ها در اطلاعات قابل استفاده است؛ یعنی اطلاعاتی که می‌تواند در افزایش درآمد، کاهش هزینه‌ها و یا هر دو آن‌ها مورداستفاده قرار گیرد.

معاملات خودپردازها حاوی الگوهای قابل توجه و روندهایی آشکار و پنهان (شامل تجمع جغرافیایی مشتریان، میزان تراکنش در هرروز، حجم تراکنش‌ها، فرهنگ منطقه و غیره) است و کشف دانش از این اطلاعات خام به دلیل حجم زیاد معاملات و الگوهای در حال تغییر همواره برای بیشتر بانکها و مؤسسات مالی یک چالش بوده است (کاماسوک، نیرندا و اونده کامبا<sup>۳</sup>، ۲۰۱۹). درواقع، خودپرداز دستگاهی است که به صورت ۲۴ ساعته برخی از عملیات بانکی را انجام می‌دهد. پیش‌بینی تعداد خودپرداز برای بانکها امری بسیار ضروری است.

پژوهش حاضر بر آن است تا به پرسش‌های زیر پاسخ دهد:  
۱- کدام شاخص‌ها در سودآوری خودپردازها تأثیرگذار است؟

۲- کدام یک از ابزارهای داده‌کاوی پیش‌بینی بهتری از افزایش امنیت خودپردازها خواهد داشت؟ اهمیت موضوع باعث شده است در مقاله حاضر با مطالعه الگوهای مختلف مبتنی بر یادگیری و در نظر گرفتن قابلیت هریک از الگوها، به شناسایی شاخص‌ها و عوامل تأثیرگذار در سودآوری و امنیت خودپردازهای بانک پرداخته شود و در نهایت مناسب‌ترین روش برای پیش‌بینی سودآوری و ضریب امنیتی خودپردازها ارائه شود.

در این مقاله، قصد داریم از الگوریتم‌های داده‌کاوی برای تخمین تعداد خودپردازهای سودده موردنیاز در هر شهر استفاده نماییم. بدین منظور، از دادگان درآمدزایی خودپردازهای کل شهر بهره می‌گیریم. در نهایت، قصد داریم بهترین مدل از لحاظ افزایش درآمدزایی و ضریب امنیتی را کشف کنیم.

## مبانی نظری

### ضریب امنیت

با ظهور فضای (دنیای) سایبری، کاربران یا همان شهروندان سایبری نیاز به احساس امنیت در فضای سایبری برای انجام امور خود دارند که این امنیت ابتدا با افزایش سطح آگاهی و دانش خود کاربران و سپس با کمک شرکت‌های امنیتی و مراجع قانونی و پلیس‌های پیشگیری از سرقت و سایبری فراهم می‌شود (صادق‌زاده و خادم‌زاده، ۱۳۸۸). ضریب امنیتی را می‌توان تابعی از عواملی مانند: کاهش احتمال سرقت یا هک؛ افزایش اقدامات لازم برای جلوگیری از سرقت، کاهش احتمالات درخطر افتادن کاربران دانست (مالیک<sup>۴</sup> و همکاران، ۲۰۲۱).

داده‌کاوی را می‌توان علمی ترکیبی از علم یادگیری ماشین، علم آمار و هوش مصنوعی دانست که با استفاده از تکنیک‌های هوشمند، دانش مطلوب را از مجموعه‌ای از داده‌ها، استخراج می‌کند. این دانش می‌تواند در قالب یک رابطه، مدل و یا ترکیبی از قواعد مشخص جهت تصمیمات آتی و یا تبیین استراتژی‌های یک سازمان مورد استفاده قرار بگیرد (نیل‌چی، مقدم، ناصر و فرهادیان، ۱۳۹۷).

سازمان‌ها معمولاً روزانه مقدار زیادی داده را در انجام عملیات تجاری خود تولید و جمع‌آوری می‌کنند. امروزه، برای پایگاه‌های داده یک شرکت عجیب نیست که مقدار داده‌های آن در حد ترابایت باشد. با این حال، علی‌رغم ثروت اطلاعاتی عظیم ذخیره‌شده حدس زده می‌شود که فقط هفت درصد کل داده‌هایی که جمع‌آوری می‌شود، مورد استفاده قرار می‌گیرد. بدین ترتیب مقدار قابل توجهی داده که بدون شک حاوی اطلاعات ارزشمند سازمانی است تا حد زیادی دست‌نخورده باقی می‌ماند. در محیط تجاری عصر اطلاعات که هر روز رقابتی‌تر می‌شود، می‌توان با استخراج اطلاعات از داده‌های استفاده‌نشده به تصمیم‌گیری‌های استراتژیک دست یافت. در طول تاریخ تحلیل داده‌ها از طریق رگرسیون و دیگر فنون آماری انجام شده است. برای استفاده از این فنون، لازم است که تحلیل‌گر مدلی خلق کند و فرآیند گردآوری دانش را سازمان دهد؛ اما امروزه این روش‌ها به‌تنهایی کافی نیستند و باید از روش‌های خودکار نیز استفاده کرد (مدهوی و همکاران، ۲۰۱۴). داده‌کاوی در

پیاده‌سازی قابلیت‌های خود از الگوریتم‌های هوش مصنوعی، مثل شبکه‌های عصبی، بهره می‌گیرد که می‌کوشند تقلیدی از عملکرد مغز انسان برای محاسبات موازی ارائه کنند. داده‌کاوی توانسته است با استفاده از شبکه‌های عصبی و دیگر مفاهیم برگرفته از هوش مصنوعی به نتایجی دست پیدا کند که حتی متخصصین این زمینه به آن‌ها دست

نیافته‌اند. این فنون امکان تحلیل مقادیر بسیار بزرگ‌تری از داده را در مقایسه با روش‌های سنتی فراهم می‌آورد. به‌علاوه، داده‌کاوی کشف و کاربرد دانش از اطلاعات خام را در پی خواهد داشت و باعث پدید آمدن پیش‌بینی‌هایی می‌شود که فراتر از عملکرد متخصصین این حوزه است (ویتن، فرانک، هال و پال، ۲۰۱۶).

### الگوریتم‌های مختلف داده‌کاوی

در ادامه، الگوریتم‌های مختلف داده‌کاوی که شامل فنون پیش‌بینی و طبقه‌بندی می‌باشند، جهت ایجاد مدل موردنظر که در این مقاله به دنبال آن هستیم، ارائه می‌شوند.

#### نزدیک‌ترین همسایه

الگوریتم  $KNN^3$  یکی از بهترین و پرکاربردترین الگوریتم‌های دسته‌بندی است که از آن استفاده گسترده‌ای در کاربردهای مختلف می‌شود. یکی از مشکلات این الگوریتم، تأثیر یکسان همه خصیصه‌ها در محاسبه فاصله رکورد جدید با همسایه‌های آن رکورد است، در صورتی که بعضی از این خصیصه‌ها برای عمل دسته‌بندی کم‌اهمیت‌ترند. این امر باعث گمراهی روند دسته‌بندی و کاهش دقت الگوریتم دسته‌بندی می‌گردد. مقایسه نتایج ارزیابی این الگوریتم با الگوریتم‌های دیگر نشان داده است که این الگوریتم می‌تواند بهبود یابد (چهاردولی، ۱۳۹۴).

#### رافست

یافتن یک واژه معادل برای عبارت Rough Sets مشکل‌ی است. در فرهنگ لغت، برای کلمه RUOGH معادل‌هایی مانند زیر، درشت، تقریبی، بی‌ادب، متلاطم و نا انطباق در نظر گرفته شده است که در میان این واژه‌ها، کلمه "تقریبی" شباهت بیشتری به مفهوم موردنظر بنیان‌گذار این

1. Madhavi  
2. Witten, Frank, Hall, & Pal  
3. k-nearest neighbors



## الگوریتم درخت طبقه‌بندی و رگرسیون<sup>۷</sup> (CART)

الگوریتم CART نخستین بار توسط اولشن<sup>۸</sup>، فریدمن<sup>۷</sup>، بریمان<sup>۸</sup> و استون<sup>۹</sup> در سال ۱۹۸۴ برای درختان رگرسیون و طبقه‌بندی طراحی شد. روش عملکرد این الگوریتم Surrogate Splitting نام دارد. این الگوریتم شامل یک متد بازگشتی است. الگوریتم CART در هر مرحله رکوردهای آموزشی را به دو زیرمجموعه تقسیم می‌کند؛ به طوری که رکوردهای هر زیرمجموعه نسبت به زیرمجموعه‌های قبلی همگن‌تر باشد. این تقسیم شدن‌ها به دفعات انجام می‌شود؛ تا شرایط توقف برقرار شود. در CART بهترین نقطه شکست با تعیین مقدار پارامتر Impurity تعیین می‌شود. اگر بهترین شکست برای یک شاخه، Impurity را از حد تعریف شده کم‌تر کند، آن انشعاب ساخته نمی‌شود. مفهوم Impurity در اینجا میزان شباهت مقدار فیلد هدف و رکوردهای رسیده به یک گره، اطلاق می‌شود. اگر ۱۰۰ درصد نمونه‌های موجود در گره، در یک دسته خاص از فیلد هدف قرار گیرد، آن گره Pure نامیده می‌شود. در این الگوریتم، یک فیلد پیشگو ممکن است به دفعات در سطوح مختلف درخت تصمیم‌گیری به کار گرفته شود. همچنین، این الگوریتم حوزه‌های طبقه‌بندی و پیش‌بینی از نوع گسسته و پیوسته را پشتیبانی می‌کند (گوترز و گومز<sup>۱۰</sup>، ۲۰۱۷).

## الگوریتم شبکه عصبی مصنوعی (ANN)

الگوریتم شبکه عصبی مصنوعی<sup>۱۱</sup> یک رویکرد قوی برای پیش‌بینی است که در آن یک شبکه عصبی به طور خاص از مجموعه‌های مختلفی از داده‌های تجربی مبتنی بر

نظریه دارد؛ اما هیچ‌یک از این کلمات بار معنایی خود واژه لاتین را ندارد. به همین علت، در این مقاله عبارت "مجموعه راف" به عنوان معادل آن به کار گرفته خواهد شد. تئوری مجموعه‌های راف در اوایل سال ۱۹۸۰ میلادی توسط پروفیسور زدیلواپاولاک<sup>۱</sup> پایه‌گذاری شد. این تئوری با تحلیل جدول‌های داده سروکار دارد. در این تئوری، جدول‌های داده می‌توانند به واسطه اندازه‌گیری یا توسط افراد متخصص و آگاه به دست آمده باشد. این دیدگاه برای بیان و بررسی مسائلی است که در آن‌ها عدم قطعیت و ابهام وجود دارد و معمولاً برای پیدا کردن ناهمگونی‌ها و ارتباطات در اطلاعات به کار می‌رود (چهاردولی، ۱۳۹۴).

## الگوریتم ماشین بردار پشتیبان<sup>۲</sup> (SVM)

الگوریتم SVM یا ماشین بردار پشتیبان یکی از روش‌های نسبتاً جدید یادگیری همراه با نظارت در مبحث یادگیری ماشین است که توسط وپنیک<sup>۳</sup> ابداع و توسعه یافت. ماشین بردار پشتیبان بر پایه تئوری آموزش آماری بنا نهاده شده است. مدل‌های ساخته شده با این الگوریتم به دلیل استفاده از اصل کمینه‌سازی ریسک ساختاری، از لحاظ تئوری عملکرد بهتری نسبت به شبکه‌های عصبی و روش‌های مشابه حین اعمال بر نمونه‌های جدید دارند. همچنین، این الگوریتم به دلیل استفاده از کرنل قابلیت بالایی را برای مدل کردن الگوهای غیرخطی و پیچیده دارد. تابع کرنل با نگاشت نمونه‌ها به فضایی که دارای ابعاد بیشتری نسبت به فضای ورودی نمونه‌های است، الگوهای پیچیده و غیرخطی را به الگوهای خطی و ساده‌تر تبدیل می‌کند. این ترفند به مدل کردن الگوهای پیچیده کمک بسیاری می‌کند (سوداهاران<sup>۴</sup>، ۲۰۱۶).

7. Friedman

8. Brieman

9. Stone

10. Gutiérrez-Esparza &amp; Gómez-Hernández

11. Artificial Neural Networks

1. Zdzisław I. Pawlak

2. Support Vector Machine

3. Vapnik

4. Suthaharan

5. Classification And Regression Tree

6. Olshen

می‌پذیرد که استفاده از روش‌های داده‌کاوی (DM) جهت پیش‌بینی موفق تماس‌های تلفنی برای فروش سپرده‌های بانکی بلندمدت، پیشنهاد شد. در (عالی‌شاهی و پورعالیزمینی، ۱۳۹۵)، به‌کارگیری روش‌های هوش تجاری و داده‌کاوی در سامانه‌های اطلاعات مدیریت در بانک مسکن پیشنهاد شد. محققین در این پژوهش با توجه به حجم انبوه اطلاعات اقدام به بررسی اطلاعات نموده و با توجه به روال‌های تعیین‌شده اقدام به استخراج قوانین، پیش‌بینی، تخمین و دسته‌بندی اطلاعات نمودند. یکی از مزیت‌های مهم سیستم هوش تجاری نسبت به گزارش‌های عادی این است که این سامانه‌ها اطلاعاتی را استخراج می‌نمایند که حتی شاید مورد توجه مدیران نبوده است.

در (اسدی، ۱۳۹۵) با استفاده از الگوریتم ژنتیک و شبکه عصبی به تشخیص پول‌شویی در سیستم بانکی پرداخته شد. در این پژوهش، روشی برای تشخیص حساب‌هایی که از آن‌ها در سیستم بانکی به‌منظور پول‌شویی استفاده می‌شود، طرح شده است. به علت وجود داده‌های زیاد در بانک‌ها، داده‌کاوی تاکنون کاربردهای زیادی در امور مالی و پولی داشته است. پس، در ابتدا با استفاده از الگوریتم ژنتیک شاخص‌ترین ویژگی‌های حساب‌های بانکی که بیشترین تأثیر در عمل پول‌شویی دارند و به‌گونه‌ای الگوهای رفتاری پول‌شویان را نشان می‌دهد، استخراج شود و سپس ویژگی‌های استخراج‌شده به همراه رکوردهای داده را به یک شبکه عصبی مصنوعی به‌منظور یادگیری و ساخت مدل تشخیص پول‌شویی ارائه می‌گردد. مدل طراحی شده با استفاده از یک مجموعه داده معتبر که توسط دانشگاه کالیفرنیا ایرواین ارائه شده است، ارزیابی و با سایر روش‌هایی که به‌منظور تشخیص در حوزه کلاهبرداری بانکی بر روی این مجموعه داده مدل‌سازی شده‌اند، مقایسه شد. از آنجایی که در فرآیندهای تشخیص مهم‌ترین معیار، دقت تشخیص است، نتایج به‌دست آمده نشان داد که این مدل از نظر صحت تشخیص، نسبت به

شبیه‌سازی مغز انسان از لحاظ عملکردهای ریاضی یاد می‌گیرد. تئوری ANN در قیاس با مغز انسان بنا نهاده شده است، که توسط عناصر بی‌شماری به نام نورون‌هایی که در لایه‌های مختلف سازمان‌یافته‌اند، تشکیل شده است. این نورون‌ها به‌هم پیوسته‌اند و بین آن‌ها اطلاعات تبادل می‌شود. هنگامی که محرک‌ها یا ورودی‌های مختلفی توسط سلول‌های عصبی دریافت می‌شوند، حالت آن‌ها تغییر پیدا کرده و اطلاعات را به سلول عصبی بعدی منتقل می‌کنند. بدین ترتیب، اطلاعات در لایه‌های مختلف نورون‌ها تا زمانی که پاسخ نهایی برای ورودی‌های اولیه به دست آید، جابه‌جا شده و تغییر پیدا می‌کنند. برای به دست آوردن پاسخ نهایی، دقیقاً مطابق با همان روشی که مغز انسان انجام می‌دهد، شبکه عصبی نیاز به یادگیری و شناخت روابط بین ورودی‌ها و خروجی‌ها دارد. بنابراین، یک ANN توسط یک‌لایه ورودی، تعدادی لایه پنهان و یک‌لایه خروجی تشکیل می‌شود که تعداد سلول‌های عصبی در هر لایه پارامتر قابل تغییر است. به‌عنوان یک رویکرد مدرن، شبکه‌های عصبی گزینه‌ای مناسب برای به دست آوردن راه‌حل طیف گسترده‌ای از مشکلات در علوم مختلف هستند که عملکرد آن پیش‌بینی و یافتن بهترین نتیجه با آموزش شبکه‌ای است. این ابزار از توانایی یادگیری عالی و قابلیت بالایی برای شناخت و مدل‌سازی روابط غیرخطی پیچیده بین متغیرهای ورودی و خروجی یک فرآیند برخوردار است (سرانو، گل‌پور و سانچز دلگادو، ۲۰۲۰).

### پیشینه پژوهش

در حوزه بانک و مسائل مربوط به آن، تحقیقات فراوانی در داخل صورت گرفته است که در ادامه به تعدادی از آن‌ها خواهیم پرداخت. در (طاهائی و آرا، ۱۳۹۵) روشی برای پیش‌بینی موفق در بازاریابی تلفنی بانک با استفاده از الگوریتم‌های داده‌کاوی ارائه شد. در این نوع بازاریابی،

مشتریان پرداختند. نتایج نشان داد که مدل ارائه شده، فرآیند شناسایی هدف را در بازاریابی مستقیم در قیاس با مدل‌های قبلی بهبود بخشیده است.

بانک‌ها مهم‌ترین نهادهای پولی در شهر هستند و رشد اقتصاد و فناوری، عرصه‌های جدیدی از خدمات برای مشتریان بانک‌ها فراهم آورده است. بنابراین، برای جذب هر چه بیشتر مشتریان در بانک‌ها، رقابت ایجاد شده و در نتیجه رضایت مشتری، تحلیل ریزش مشتریان، بهبود ارائه خدمات و پشتیبانی به مشتریان، پیگیری سوءاستفاده‌های مالی، کشف تقلب و پول‌شویی، محاسبه درآمد و زیان وام‌ها، تحلیل و رتبه‌بندی شعب بر اساس کارایی عملکردشان، تجهیز مدیران شعب به ابزار تحلیلی برای پیمایش در کلیه داده‌ها از بیرونی‌ترین سطح، به سطوح درونی اهمیت خاصی پیدا کرده است. از این رو، در (کریمی، ۱۳۹۵) به بررسی مفهوم هوش تجاری و حوزه‌های دربرگیرنده آن در بانک‌ها، مدیریت دانش، انبار داده، داده‌کاری، نحوه یکپارچه نمودن داده‌ها و تبدیل آن‌ها به اطلاعات مفید جهت تصمیم‌گیری مؤثر پرداخته شده است.

سایر روش‌های مدل‌سازی جهت تشخیص بر روی این مجموعه داده با ۱۶/۰۹ درصد دارای عملکرد بالاتری است. در (محمدی و علیوند، ۱۳۹۵) از روش‌های داده‌کاوی برای ارزیابی و اعتبارسنجی مشتریان استفاده شده است. هدف این پژوهش در نهایت ارائه یک مدل برای اعتبارسنجی مشتریان تسهیلات قبل از اعطای تسهیلات است تا از ایجاد مطالبات معوق و سوخت تسهیلات جلوگیری شود.

در (بازگیر، ۱۳۹۵) ادعا شد که در شهرهای مختلف تجهیز منابع در نظام بانکی ایران با طبقه‌بندی مشتریان با استفاده از الگوریتم‌های داده‌کاوی و شبکه‌های عصبی مصنوعی، بانک‌ها را قادر به حفظ وفاداری مشتریان قدیم و جذب مشتریان جدید خواهد کرد. یکی از روش‌های داده‌کاوی، درخت تصمیم‌گیری است. چنانچه درخت تصمیم‌گیری مناسب و درستی اتخاذ شود، می‌توان طبقه‌بندی مشتریان را به‌طور بهینه انجام داد. برای دسته‌بندی مشتریان در حوزه‌های مختلف، بخصوص در حوزه درگاه‌های الکترونیکی، درخت تصمیم‌گیری نهایی بر اساس بهینگی و دقت بوده و مشتریان را به سه دسته بالا، متوسط و پایین طبقه‌بندی می‌کند و بیشترین مشتریان جدیدی را که متقاضی استفاده از خدمات هستند، پیدا می‌کند. در این پژوهش ضمن بررسی شیوه‌های مختلف تجهیز منابع طبقه‌بندی گروه‌های مختلف مشتریان در سیستم بانکی با استفاده از پارامترها و شاخص‌هایی نظیر ارزش طول عمر مشتری در مدیریت ارتباط مشتری و برای الگوسازی آن از شبکه‌های عصبی استفاده شده است.

در (خطیر، بائی، و خنداخند، ۱۳۹۵) مدل ترکیبی جدیدی برای دسته‌بندی مشتریان و شناسایی مشتریان هدف ارائه می‌گردد که بر روی یک مجموعه داده بانک آزمایش شده و نتایج ارائه شده است. ابتدا با استفاده از مطالعات کتابخانه‌ای، مصاحبه با خبرگان و رجوع به پایگاه داده اینترنتی بانک، شاخص‌های دسته‌بندی مشتریان بانک استخراج و با استفاده از الگوریتم ژنتیک انتخاب شد و سپس با استفاده از روش درخت تصمیم به دسته‌بندی



جدول ۱- مقالات ارائه‌شده در سال‌های اخیر

ردیف	سال	نام مقاله	نویسنده	الگوریتم استفاده شده
<b>پژوهش‌های داخلی</b>				
۱	۱۳۹۷	پیش‌بینی ریسک تسهیلات پرداختی با استفاده از ابزارهای داده‌کاوی	(نیل‌چی و همکاران، ۱۳۹۷)	درخت تصمیم، شبکه عصبی و ماشین بردار پشتیبان
۲	۱۳۹۷	پیش‌بینی رویگردانی مشتریان بانک با استفاده از روش داده‌کاوی	(نجمی، راد و شعار، ۲۰۱۸)	ماشین بردار پشتیبان و درخت تصمیم
۳	۱۳۹۵	پیش‌بینی رفتار مشتریان با استفاده از تکنیک شبکه‌های عصبی مصنوعی به‌کارگیری روش‌های هوش تجاری و داده‌کاوی در سیستم‌های اطلاعات	(ملائی و پارسا، ۱۳۹۵)	شبکه عصبی مصنوعی
۴	۱۳۹۵	مدیریت مطالعه موردی: پیاده‌سازی مفاهیم و کاربردهای هوش تجاری در بانک مسکن	(عالیشاهی و قربان پورعالیزمینی، ۱۳۹۵)	شبکه عصبی مصنوعی
۵	۱۳۹۵	تشخیص پول‌شویی در سیستم بانکی با استفاده از الگوریتم ژنتیک و شبکه عصبی	(اسدی، ۱۳۹۵)	شبکه عصبی مصنوعی
۶	۱۳۹۵	مطالعه جامع اعتبارسنجی مشتریان بانک و ارائه یک روش اعتبارسنجی با تکنیک داده‌کاوی و منطق فازی	(فردین ابدالی و مصطفی، ۱۳۹۵)	درخت تصمیم فازی
۷	۱۳۹۵	دسته‌بندی مشتریان برای به‌کارگیری بازاریابی مستقیم از طریق داده‌کاوی (مطالعه موردی: بانک کشاورزی) طبقه‌بندی متقاضیان تسهیلات	(اکبری خطیر، همایون‌زاده‌بائی و خنداخند، ۱۳۹۵)	درخت تصمیم
۸	۱۳۹۱	اعتباری بانکی با استفاده از داده‌کاوی و منطق فازی	(تقوی‌فرد و نادعلی، ۲۰۱۲)	درخت تصمیم فازی
۹	۱۳۹۱	بررسی راهکارهای امنیتی در ماشین‌های ATM، کارت‌خوان و آسیب‌شناسی آن‌ها	(اصغری و توانگر، ۱۳۹۱)	الگوریتم بیومتریک
<b>پژوهش‌های خارجی</b>				
۱۰	۲۰۱۹	پیش‌بینی رفتار مشتریان بانک با استفاده از داده‌کاوی	(ژو <sup>۱</sup> و همکاران، ۲۰۱۹)	شبکه عصبی مصنوعی

ادامه جدول ۱- مقالات ارائه‌شده در سال‌های اخیر

ردیف	سال	نام مقاله	نویسنده	الگوریتم استفاده شده
۱۱	۲۰۱۹	کاربرد روش‌های داده‌کاوی برای تجزیه و تحلیل سودآوری مشتری و رفتار خرید مشتریان	(گودینهو، دیاس و تورس، ۲۰۱۹)	ترکیبی
۱۲	۲۰۱۹	طراحی مدل پیش‌بینی خطرات وام بانکی با استفاده از روش‌های داده‌کاوی مدل ارزیابی ریسک اعتباری	(کاور، بهدادال و سینگ، ۲۰۱۹)	درخت طبقه‌بندی
۱۳	۲۰۱۸	دومرحله‌ای برای برنامه‌های خرده‌فروشی وام بانکی با استفاده از تکنیک داده‌کاوی درخت تصمیم پیش‌بینی مسائل امنیتی محاسبات ابری	(سوداکار و ردی، <sup>۳</sup> ۲۰۱۹)	درخت طبقه‌بندی
۱۴	۲۰۱۷	با استفاده از الگوریتم‌های شبکه عصبی مصنوعی (ANNS) در سازمان‌های بانکی رویکرد داده‌کاوی برای بازاریابی مستقیم محصولات بانکی با تجزیه و تحلیل سود / هزینه	(سوداکار و ردی، ۲۰۱۹)	شبکه عصبی مصنوعی
۱۵	۲۰۱۷	طراحی الگویی برای بهبود سیستم‌های بانکی بر اساس پیش‌بینی منافع مشتری: با استفاده از تکنیک‌های داده‌کاوی تجزیه و تحلیل خواسته مشتری در بخش بانکی با استفاده از تکنیک‌های داده‌کاوی	(میتیک و همکاران، <sup>۴</sup> ۲۰۱۷)	درخت طبقه‌بندی
۱۶	۲۰۱۶	طراحی الگویی برای بهبود سیستم‌های بانکی بر اساس پیش‌بینی منافع مشتری: با استفاده از تکنیک‌های داده‌کاوی تجزیه و تحلیل خواسته مشتری در بخش بانکی با استفاده از تکنیک‌های داده‌کاوی	(مظهری‌تژاد، ذوالفقارزاده، خدنگی و سعدآبادی، ۲۰۱۶)	نزدیک‌ترین همسایه، ماشین بردار پشتیبان
۱۷	۲۰۱۵	تجزیه و تحلیل خواسته مشتری در بخش بانکی با استفاده از تکنیک‌های داده‌کاوی تجزیه و تحلیل خدمات دستگاه خودپرداز با پیش‌بینی از طریق داده‌کاوی	(اویینی و آدیمو، <sup>۵</sup> ۲۰۱۵)	کی‌مینز <sup>۶</sup> (هوش مصنوعی)
۱۸	۲۰۱۴	تجزیه و تحلیل خدمات دستگاه خودپرداز با پیش‌بینی از طریق داده‌کاوی	(مدهاوی، ۲۰۱۴)	درخت طبقه‌بندی

1. Godinho, Dias, & Torres  
 2. Kaur, Bhaddal, & Singh  
 3. Sudhakar & Reddy  
 4. Mitik et al.  
 5. Oyeniyi, Adeyemo  
 6. K-Means

## روش‌شناسی پژوهش

ماهیت این تحقیق استفاده از داده‌کاوی به‌منظور پیش‌بینی تعداد خودپردازهای سودآور است که با توجه به آن، می‌توان ماهیت پژوهش را داده‌محور دانست. تحقیق حاضر بر مبنای کشف دانش از پایگاه داده‌های بانک تجارت شهر همدان نهاده شده و از شهر دارد جهانی CRISP-DM برای انجام تحقیق بهره گرفته شده است که مراحل آن شامل درک مسأله کسب‌وکار، درک داده‌ها، آماده‌سازی داده‌ها، الگوسازی، ارزیابی نتایج و به‌کارگیری الگو به انضمام ارتباط موجود بین آن‌ها در شکل ۱ نمایش داده می‌شود.



شکل ۱- فرایند داده‌کاوی توسط شهر دارد Crisp-DM (ورث و هپ، ۲۰۰۰)

## درک داده‌ها

در این بخش از پژوهش، مراحل زیر طی شده است:

### برگزیدن معیارهای اساسی در کارکرد خودپردازها

به‌منظور برگزیدن این معیارها که نمایانگر کارکرد یک خودپرداز هستند، از خبرگان امور بانکی فعال در بخش خودپردازها تحقیق به عمل آمد و در نهایت با استفاده از همین نظرات، هر خودپرداز دارای پنج متغیر اساسی شامل مواردی همچون نوع مالکیت، مدل خودپرداز، تعداد تراکنش در ماه، مجموع مبلغ تراکنش‌ها به ریال و مجموع کارمزدها به ریال در ماه است (جدول ۲)، شناسایی شد. با توجه به توضیحات قبلی و همان‌طور که در جدول ۲ شاهدیم، خودپرداز شماره یک زیان‌ده به حساب می‌آید؛ زیرا کارمزد آن کم‌تر از دو میلیون تومان است؛ ولی خودپرداز دوم سودآور بوده و ادامه سرویس‌دهی آن برای بانک مفید است. در این تحقیق، تعداد شعب بانکی یک شهر تقسیم بر تعداد کل شعب، تعداد خودپردازهای هر شهر تقسیم بر تعداد کل در شهر، تعداد خودپردازهای مشارکتی تقسیم بر تعداد کل در یک شهر، تعداد مدل سالنی تقسیم بر تعداد کل در یک شهر، کل تراکنش‌های خودپردازها تقسیم بر کل تراکنش‌ها و کل کارمزد حاصل از خودپردازهای شهر تقسیم بر کل کارمزدها و تعداد

## درک مسأله کسب‌وکار

در این بخش از فرایند تحقیق، ابتدا باید اهداف اصلی کسب‌وکار تعیین شود. در این پژوهش، اصلی‌ترین هدف کسب‌وکار یا همان بانک‌های مورد مطالعه پیش‌بینی تعداد خودپردازهای سودآور در آینده است تا برحسب آن بتوانیم برای سال‌های آتی برنامه‌ریزی داشته باشیم و این کار بر اساس کشف دانش از آمارهای موجود از خودپردازهایی

مهم بر عهده الگو گذاشته می شود و محققین دخالتی در این امر نداشتند.

خودپردازهای سودآور هر خودپردازهای سودآور ماه قبل را به عنوان ورودی و تعداد خودپردازهای سودآور ماه بعد به عنوان خروجی استفاده می نماییم؛ اما تشخیص متغیرهای

جدول ۲- نمونه ای از اطلاعات خودپرداز (منبع: بانک تجارت شهر همدان، ۱۳۹۸).

شماره	نوع مالکیت	مدل	تعداد تراکنش	تراکنش (ریال)	کارمزد (ریال)	تعداد نگهداری	تراکم جمعیت (km <sup>2</sup> )	مسافت تا کلاتری (km)	تعداد دوربین های منطقه
۱	مشارکتی	سالنی	۱۵۲	۳۰۰,۲۶۴,۰۰۰	۸۳۹,۱۷۵	۲	۳۷	۱/۷	۱۷۴
۲	مشارکتی	دیواری	۵۰۷۴	۱۲,۱۴۴,۴۴۲,۲۲۹	۲۰,۱۳۸,۵۷۵	۱	۴۲	۳/۴	۲۰

### جمع آوری داده ها

با ادغام برخی شاخص های مؤثر در پیش بینی تعداد خودپردازها، نهایتاً ۵ شاخص جهت الگوسازی انتخاب شد. از بین شاخص های منتخب، تنها تقریباً ۹۰ درصد شاخص ها دارای اطلاعات کامل در ۳۳ توالی بودند. به منظور افزایش کیفیت داده های جمع آوری شده، آماده سازی داده ها از طریق پالایش آن ها از داده های ناقص، پرت، و ناسازگار، تلخیص، یکپارچه سازی، و درون یابی برخی داده های مفقود و غیره صورت پذیرفت. به این منظور، از الگوریتم شبکه عصبی مصنوعی استفاده شد.

جامعه آماری این پژوهش آمارهای مستخرج از خودپردازهای بانک تجارت شهر همدان از ابتدای سال ۱۳۹۵ تا پایان سال ۱۳۹۸ است که شامل اطلاعات خودپردازهای تمام بانک های تجارت مناطق چهارگانه شهری همدان در همراه بود. آمار این داده ها از طریق اطلاعات موجود در نظام بانکداری متمرکز بانک و مکاتبه با شعب جمع آوری گردید.

### ساختار و آماده سازی داده ها

از بین توالی ۴۸ ماهه که در بخش قبل عنوان شد، تعداد ۱۵ ماه به دلیل موجود نبودن آمار برخی بانک ها از پژوهش کنار گذاشته شد. به منظور افزایش کیفیت داده های جمع آوری شده بنا بر برخی دلایل از جمله عدم ثبت، ورود ناقص، یا اشتباه برخی داده های خودپردازها در سامانه و نبود یا وجود تناقض در داده های مربوط به شاخص های اشاره شده، فرایند آماده سازی داده ها با عاری ساختن آن ها از داده های ناقص، مغشوش و ناسازگار، تلخیص و یکپارچه سازی و درون یابی برخی داده های مفقود با الگوریتم شبکه عصبی مصنوعی صورت پذیرفت. از بین ۳۳ توالی باقیمانده ۸۰٪ آن ها دارای اطلاعات کامل همه شاخص ها بود.

### الگوسازی

در گام بعدی، پس از آماده سازی داده ها به منظور پیش بینی تعداد خودپردازهای سودده، از الگوریتم های رافست، نزدیک ترین همسایه، درخت تصمیم، ماشین بردار پشتیبان و شبکه عصبی در محیط نرم افزار متلب<sup>۱</sup> استفاده شد. برای سنجش کارایی الگوریتم ها پس از مرحله تعلیم، مدل های آموزش دیده را با دادگان آزمون مورد آزمایش قرار می دهیم.

خودپردازهای شهر تقسیم بر کل کارمزدها و تعداد خودپردازهای سودآور هر شهر تقسیم بر کل تعداد خودپردازهای سودآور ماه قبل ورودی‌های الگوریتم و خروجی تعداد خودپردازهای سودآور ماه آخر است. در این تحقیق، در روش رافست، از سه الگوریتم ژنتیک، جانسون و هولتس استفاده شده است که خطای تخمین هریک از روش‌ها را در جدول ۳ مشاهده می‌کنیم.

جدول ۳- مقایسه  $RMSE$  روش رافست.

نام روش	ژنتیک	جانسون	هولتس
RMSE	۰/۵۱۴۲	۰/۶۴۸۲	۰/۷۸۴۲

در روش نزدیک‌ترین همسایه از پنج همسایه استفاده شد و در شبکه عصبی از ده نرون لایه پنهان بهره بردیم. در جدول ۳ خطای تمام روش‌ها باهم مقایسه شده است.

جدول ۴- مقایسه  $RMSE$  مدل‌ها.

نام روش	رافست (ژنتیک)	نزدیک‌ترین همسایه	درخت تصمیم	SVM	شبکه عصبی
RMSE داده آموزش	۰/۵۱۴۲	۰/۴۲۱۴	۰/۲۱۵۰	۰/۰۷۹۸	۰/۰۹۳۲
RMSE داده آزمون	۰/۵۳۲۵	۰/۴۶۲۱	۰/۲۴۱۵	۰/۰۸۰۹	۰/۱۱۲۳

در ادامه، برای رتبه‌بندی ورودی‌ها، متغیرهای مؤثر در نتایج حاصل از الگوریتم درخت تصمیم به شرح جدول ۵ ارائه شده است:

برای این منظور، داده ماه آخر و یک ماه قبل از آن را به‌عنوان داده‌های آزمایش (تعداد خودپرداز ماه آخر خروجی مطلوب و داده‌های ماه قبل ورودی الگوریتم) و داده‌های ماه‌های قبلی را به‌عنوان دادگان آزمایش در نظر می‌گیریم و جهت ارزیابی روش‌ها، معیار جذر میانگین مربعات خطا<sup>۱</sup> به‌کار می‌رود:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

که در آن  $\hat{y}_i$  مقدار تخمینی،  $y_i$  مقدار واقعی و  $n$  تعداد تمام نمونه‌هاست.

با توجه به موارد عنوان‌شده، تعداد شعب بانکی یک شهر تقسیم بر تعداد کل شعب، تعداد خودپردازهای هر شهر تقسیم بر تعداد کل در شهر، کل تراکنش‌های خودپردازها تقسیم بر کل تراکنش‌ها، کل کارمزد حاصل از

با توجه به نتایج به این نتیجه می‌رسیم که "ماشین بردار پشتیبان" بهترین روش در پیش‌بینی تعداد خودپردازهای سودده و امن است.



جدول ۵- اندازه اهمیت شاخص‌ها برای هر منطقه شهر همدان ( برگرفته از الگوریتم درخت تصمیم)

اهمیت	شاخص	ردیف
۰/۰۷	تعداد شعب بانکی یک منطقه تقسیم بر تعداد کل شعب	۱
۰/۰۸	تعداد خودپردازهای هر منطقه تقسیم بر تعداد کل در منطقه	۲
۰/۳۸	کل تراکنش‌های خودپردازها تقسیم بر کل تراکنش‌ها	۳
۰/۴۳	کل کارمزد حاصله از خودپردازهای منطقه تقسیم بر کل کارمزدها	۴
۰/۶۱	تعداد خودپردازهای سودآور هر منطقه تقسیم بر کل تعداد خودپردازهای سودآور	۵
۰/۲	تراکم جمعیت	۶
۰/۳۲	میزان سرقت در منطقه	۷
۰/۴۵	مسافت تا نزدیک‌ترین کلانتری	۸
۰/۱۵	تعداد دوربین‌های منطقه	۹
۰/۳۶	تعداد نگهبان	۱۰

## ارزیابی نتایج

**تعداد شعب بانکی یک منطقه:** هرچند از بین سایر شاخص‌ها، کم‌ترین ارزش را دارد، همچنان این شاخص می‌تواند برای تخمین تعداد خودپرداز سودده ماه‌های آتی مفید باشد.

**تراکم جمعیت:** عبارت است از تقسیم شماره افراد ساکن در یک منطقه بر مساحت آن منطقه.

**میزان سرقت در منطقه:** بر اساس تعداد سرقت‌هایی که از پلیس آگاهی به‌دست آمده است، تعریف می‌شود.

**مسافت تا نزدیک‌ترین کلانتری:** بر اساس فرمول (۲) محاسبه می‌شود:

$$Distance = \sqrt{(Y_b - Y_a)^2 + (X_b - X_a)^2}$$

فرمول (۲)

Distance = مسافت تا نزدیک‌ترین کلانتری

$(X_a, Y_a)$  = مختصات نزدیک‌ترین کلانتری

$(X_b, Y_b)$  = مختصات محل نصب خودپرداز

لذا، توجه به این شاخص‌ها قبل از تخصیص خودپرداز و تعیین تعداد موردنیاز در تصمیم‌گیری و پیش‌بینی بسیار

مطابق نتایج حاصل از الگوریتم‌ها، مهم‌ترین متغیر مدت قرارداد با ضریب ۰.۶۱ است که می‌توان گفت تعداد خودپردازهای سودده ماه آتی باید متناسب با تعداد خودپردازهای سودده ماه فعلی باشد تا احتمال ضرردهی کاهش یابد. به‌عبارت‌دیگر، یکی از معضلات اساسی در تخصیص تعداد خودپرداز اضافی و بی‌کاربرد بی‌توجهی به این اصل کلیدی است. در ادامه، به توضیح هر یک از معیارها پرداخته شده است:

**کل کارمزد:** از لحاظ نظری، کارمزد بالاتر توسط خودپردازهای نصب‌شده فعلی ارتباط مستقیم بر تعداد خودپرداز سودده در ماه آتی دارد.

**کل تراکنش‌ها:** این موضوع نیز روشن است؛ زیرا تعداد تراکنش بیشتر احتمال بالاتر بودن کارمزد را افزایش و نیاز به خودپردازهای بیشتر در راستای تولید سود بیشتر را ارتقا می‌دهد. **تعداد خودپردازهای هر منطقه:** این شاخص از شاخص‌های قبلی کم‌ارزش‌تر است ولی همچنان اهمیتی هرچند کم بر روی تعداد خودپرداز سودده ماه آتی دارد.

همدان به تفکیک در جدول ۶ نشان داده شده است. همچنین، این الگوریتم یک‌بار برای معیارهای یک تا پنج جدول اجرا شد و تعداد خودپردازهای زیان ده هر منطقه شناسایی شد و سپس مجدداً با به‌کارگیری معیارهای شش تا ده جدول ۵ در این الگوریتم، خودپردازهای نا امن نیز مشخص شد و در جدول ۶، نتایج آن مشخص شد.

حائز اهمیت است. این شاخص‌ها مبنایی برای پیش‌بینی تعداد خودپرداز سودده خواهند بود. همچنین، الگوریتم ماشین بردار پشتیبان با جذر میانگین مربعات خطای ۸ درصد، قابلیت برآورد بهتری را نسبت به سایر الگوریتم‌ها دارد و نشان دهنده این است که بهترین برآورد از نظر سوددهی و امنیت جانمایی و تعداد بهینه خودپردازهای بانک تجارت در مناطق شهری همدان دارد. یافته‌های حاصل از این الگوریتم در خصوص مناطق ۴ گانه شهری

جدول ۶. مکان‌یابی تعداد خودپردازهای سودده امن براساس مناطق ۴ گانه شهر همدان

منطقه شهری	تعداد خودپرداز فعلی	تعداد خودپرداز پیشنهاد شده پژوهش	تعداد خودپرداز زیان دیده شناسایی شده	تعداد خودپرداز نا امن شناسایی شده
۱	۳۵	۳۱	۲	۲
۲	۲۶	۲۰	۵	۱
۳	۲۴	۱۸	۲	۴
۴	۳۹	۲۹	۱	۹
جمع	۱۲۴	۹۸	۱۰	۱۶

### نتیجه‌گیری

امنیت خودپردازها از آنجایی حائز اهمیت است که سهم بسیار مهمی در تراکنش‌های مالی بانک‌ها به عهده دارد و با زندگی مالی مردم جامعه عجین شده است و جزئی از مهم‌ترین رفتارهای مالی سپرده‌گزاران و مشتریان بانک است و بنابراین، مهم است که امنیت آن‌ها دائماً رصد گردد و در بهینه‌ترین حالت ممکن خود از نظر مکانی نصب گردد تا هم سوددهی بالاتری داشته باشد و هم از سرقت، جلوگیری به عمل آورد و نرخ سرقت را به کم‌ترین حد خود برساند. این پژوهش توانست با ارائه یک مدل بهینه با استفاده از پارامترهای تشخیص داده شده از درخت تصمیم، و اجرای آن توسط الگوریتم ماشین بردار پشتیبان داده‌کاوی، تعداد صحیح خودپردازهای بانک تجارت را هم از حیث سوددهی و هم از حیث امنیت قابل قبول پلیس، ارائه دهد

اجرای روش‌ها / راهکارهای پیش‌بینی تعداد، به بانک‌ها اجازه می‌دهد از اهداف موجود برای اهداف سرمایه‌گذاری کامل استفاده کنند. اساساً، اجرای مدل‌های بهینه‌سازی تعداد دستگاه نه تنها متمرکز کردن فرآیند تدارکات مدیریت هزینه برای هر شعبه و دستگاه خودپرداز بلکه به بانک‌ها نیز امکان می‌دهد تا کل هزینه‌های متحمل شده را به حداقل برسانند. با انجام عملیات نقدی و افزایش حداقل میزان سود نقدی رضایت مشتری را افزایش می‌دهد. بر این اساس، پژوهش حاضر به پیش‌بینی تعداد خودپردازهای سودآور و امن در شعب بانک با استفاده از الگوریتم‌های داده‌کاوی پرداخت. نتایج این پژوهش نشان داد که ماشین بردار پشتیبان بهترین روش در پیش‌بینی تعداد خودپردازها است و می‌توان از آن جهت پیش‌بینی تعداد خودپردازهای هر شهر استفاده کرد.

تا مدیران بانک بتوانند تخصیص بهینه‌ای برای تعداد خودپردازها متناسب با جانمایی آن‌ها در مناطق ۴ گانه شهری داشته باشند و در این زمینه نیز پلیس پیشگیری می‌تواند با اجرای این مدل بر روی سایر بانک‌ها، خودپردازهای ناامن را شناسایی کرده و به بانک هشدارهای لازم جهت جمع‌آوری آن و یا تغییر مکان آن‌ها را ارائه کند.

## تشکر و قدردانی

بدین وسیله از کلیه پرسنل محترم فرماندهی نیروی انتظامی، به‌خصوص پلیس پیشگیری ناجای همدان که نهایت همکاری را با نویسندگان این پژوهش داشتند، کمال تشکر و قدردانی را داریم. همچنین، از مدیران بانک تجارت همدان نیز که جهت جمع‌آوری اطلاعات کمک شایانی کردند، سپاسگزاریم.

## منابع

- اسدی، م. (۱۳۹۵). تشخیص پول‌شویی در سیستم بانکی با استفاده از الگوریتم ژنتیک و شبکه عصبی. کنفرانس بین‌المللی پژوهش‌های نوین در علوم مهندسی، تهران.
- اکبری‌خطیر، م؛ همایون‌زاده‌بائی، ر؛ و خنداخند، م. (۱۳۹۵). دسته‌بندی مشتریان برای به‌کارگیری بازاریابی مستقیم از طریق داده‌کاوی (مطالعه موردی: بانک کشاورزی). کنفرانس جهانی مدیریت، اقتصاد حسابداری و علوم انسانی در آغاز هزاره سوم.
- اصغری، علی و توانگر، مولود. (۱۳۹۱). بررسی راهکارهای امنیتی در ماشین‌های ATM، کارت‌خوان و آسیب‌شناسی آن‌ها. اولین همایش منطقه‌ای پژوهش‌ها و راهکارهای نوین در حسابداری و مدیریت، تنکابن.
- بازگیر، ع. (۱۳۹۵). واکاوی روش‌های تجهیز و تخصیص منابع در نظام بانکی ایران با طبقه‌بندی مشتریان به‌وسیله الگوریتم‌های داده‌کاوی و شبکه‌های عصبی مصنوعی.
- دومین کنفرانس ملی رویکردهای نوین در مهندسی کامپیوتر و برق، رودسر.
- بانک مرکزی. (۱۳۸۱). مقررات حاکم بر مرکز شتاب. بانک مرکزی جمهوری اسلامی ایران.

- تقوی‌فرد، م؛ نادعلی، ا. (۲۰۱۲). طبقه‌بندی متقاضیان تسهیلات اعتباری بانکی با استفاده از داده‌کاوی و منطق فازی. مطالعات مدیریت صنعتی، ۹(۲۵)، ۸۵-۱۰۷.
- چهاردولی، م. (۱۳۹۴). بررسی کارایی روش رافت و نزدیک‌ترین همسایه در تخمین قیمت سهام پذیرفته‌شده در بورس اوراق بهادار تهران (پایان‌نامه کارشناسی ارشد). دانشگاه آزاد اسلامی، همدان.
- خطیر، م. ا؛ بائی، ر. و خنداخند، م. (۱۳۹۵). دسته‌بندی مشتریان برای به‌کارگیری بازاریابی مستقیم از طریق داده‌کاوی (مطالعه موردی: بانک کشاورزی). کنفرانس جهانی مدیریت، اقتصاد حسابداری و علوم انسانی در آغاز هزاره سوم، شیراز.
- صادقی، ت. (۱۳۹۱). شاخص‌های پذیرش و توسعه بانکداری الکترونیک در ایران (با نگاهی به تجارب سایر شهرها). نشریه پژوهش‌های بولی و بانکی، ۳(۸)، ۲۴۳-۲۷۷.
- صادق‌زاده، مهدی؛ خادم‌زاده، احمد. (۱۳۸۸). ارائه طرحی جهت افزایش ضریب امنیت در دسترسی کاربران به اطلاعات نظام سلامت. کنفرانس شهر الکترونیک، پژوهشگاه فناوری اطلاعات و ارتباطات جهاد دانشگاهی، دوره ۲.
- طاهانی، س. ل. و آرا، ف. ص. (۱۳۹۵). ارائه روشی برای پیش‌بینی موفق در بازاریابی تلفنی بانک با استفاده از الگوریتم‌های داده‌کاوی. اولین کنفرانس بین‌المللی دستاوردهای نوین پژوهشی در مهندسی برق و کامپیوتر.
- عالیشاهی، م؛ و پورعالی‌زمینی، ف. ق. (۱۳۹۵). به‌کارگیری روش‌های هوش تجاری و داده‌کاوی در سیستم‌های اطلاعات مدیریت مطالعه موردی: پیاده‌سازی مفاهیم و کاربردهای هوش تجاری در بانک مسکن. نخستین کنفرانس بین‌المللی پارادیم‌های نوین مدیریت هوشمندی تجاری و سازمانی، تهران.
- عالیشاهی، م؛ قربان پورعالی‌زمینی، ف. (۱۳۹۵). به‌کارگیری روش‌های هوش تجاری و داده‌کاوی در سیستم‌های اطلاعات مدیریت مطالعه موردی: پیاده‌سازی مفاهیم و کاربردهای هوش تجاری در بانک مسکن. نخستین کنفرانس بین‌المللی پارادیم‌های نوین مدیریت هوشمندی تجاری و سازمانی.

25. Gutiérrez-Esparza, J. C., & Gómez-Hernández, J. J. (2017). Inverse Modeling Aided by the Classification and Regression Tree (CART) Algorithm. In *Geostatistics Valencia 2016* (pp. 805-819). Springer, Cham.
26. Kamusweke, K., Nyirenda, M., & onde Kabemba, M. (2019). A Data Mining Model for Predicting and Forecasting Fraud in Banks."
27. Kaur, M., Bhaddal, P., & Singh, G. (2019). Calculation of client credit risk prediction in banking sector using data mining.
28. Madhavi, S., Abirami, S., Bharathi, C., Ekambaram, B., Sankar, T. K., Nattudurai, A., & Vijayarangan, N. (2014). ATM Service Analysis Using Predictive Data Mining. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 8(2).
29. Mallik, S., Halder, S., Saha, P., & Mukherjee, S. (2021). Multi-factor Authentication-Based E-Exam Management System (EEMS). In *Proceedings of International Conference on Frontiers in Computing and Systems* (pp. 711-720). Springer, Singapore.
30. Mitik, M., Korkmaz, O., Karagoz, P., Toroslu, I. H., & Yucel, F. (2017). Data mining approach for direct marketing of banking products with profit/cost analysis. *The Review of Socionetwork Strategies*, 11(1), 17-31.
31. Motaharnejad, M. S., Zolfagharzadeh, M. M., Khadangi, E., & Sadabadi, A. A. (2016). Designing a Model for Improving Banking Recommender Systems Based on Predicting Customers' Interests: Application of Data Mining Techniques. *Journal of Information Technology Management*, 8(2), 393-314.
32. Oyeniyi, A. O., Adeyemo, A. B., Oyeniyi, A. O., & Adeyemo, A. B. (2015). Customer churn analysis in banking sector using data mining techniques. *Afr J Comput ICT*, 8(3), 165-174.
33. Rahman, M. M., & Saha, A. R. (2019). A comparative study and performance analysis of ATM card fraud detection techniques. *Journal of Information Security*, 10(03), 188.
۱۵. فردین ابدالی، م؛ مصطفی، ع. (۱۳۹۵). مطالعه جامع اعتبارسنجی مشتریان بانک و ارائه یک روش اعتبارسنجی با تکنیک داده‌کاوی و منطق فازی. *اولین کنفرانس ملی رویکردهای نو در مهندسی برق و کامپیوتر*.
۱۶. کریمی، س. (۱۳۹۵). تأثیر هوش تجاری بر صنعت بانکداری. *اولین کنفرانس بین‌المللی هوش تجاری ایران*، تهران.
۱۷. کیهان. (۱۳۵۱). ورود ATM به ایران. *روزنامه کیهان*، ص ۱۹.
۱۸. محمدی، ف.؛ اعلیوند، م. (۱۳۹۵). مطالعه جامع اعتبارسنجی مشتریان بانک و ارائه یک روش اعتبارسنجی با تکنیک داده‌کاوی و منطق فازی. *اولین کنفرانس ملی رویکردهای نو در مهندسی برق و کامپیوتر*، لرستان، خرم‌آباد.
۱۹. ملائی، م؛ پارسا، س. (۱۳۹۵). پیش‌بینی رفتار مشتریان با استفاده از تکنیک شبکه‌های عصبی مصنوعی. *شبک، ۲* (۳).
۲۰. نجمی، پ؛ راد، ع؛ شعار، م. (۲۰۱۸). پیش‌بینی رویگردانی مشتریان بانک با استفاده از روش داده‌کاوی. *فصلنامه مدیریت صنعتی*، ۱۳(۴۴)، ۹۹-۱۱۱.
۲۱. نوری‌بورجردی، پ؛ جلیلی، م؛ مردانی، ف. (۱۳۸۹). بررسی تأثیر تمرکز و سایر عوامل در صنعت بانکداری بر سودآوری بانک‌های دولتی. *پژوهش‌های پولی-بانکی*، ۱(۱)، ۱۷۵-۲۰۲.
۲۲. نیلچی، م؛ مقدم، خ؛ ناصر، ع؛ فرهادیان، ع. (۱۳۹۷). پیش‌بینی ریسک تسهیلات پرداختی با استفاده از ابزارهای داده‌کاوی. *فصلنامه پژوهش‌های پولی-بانکی*، ۱۱(۳۸)، ۶۲۵-۶۵۴.
23. Bilir, C., & Döşeyen, A. (2018). Optimization of ATM and branch cash operations using an integrated cash requirement forecasting and cash optimization model. *Business & Management Studies: An International Journal*, 6(1), 237-255.
24. Godinho, P., Dias, J., & Torres, P. (2019). An Application of Data Mining Methods to the Analysis of Bank Customer Profitability and Buying Behavior. *Data Analysis and Applications 1: Clustering and Regression, Modeling-estimating, Forecasting and Data Mining*, 2, 225-240.

34. Serrano, D., Golpour, I., & Sánchez-Delgado, S. (2020). Predicting the effect of bed materials in bubbling fluidized bed gasification using artificial neural networks (ANNs) modeling approach. *Fuel*, 266, 117021.
35. Sudhakar, M., & Reddy, C. V. K. (2016). Two step credit risk assessment model for retail bank loan applications using Decision Tree data mining technique. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 5(3), 705-718.
36. Suthaharan, S. (2016). Machine learning models and algorithms for big data classification. *Integr. Ser. Inf. Syst*, 36, 1-12.
37. Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (Vol. 1). London, UK: Springer-Verlag.
38. Witten, I. H., & Frank, E. (2002). Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record*, 31(1), 76-77.
39. Zhou, X., Bargshady, G., Abdar, M., Tao, X., Gururajan, R., & Chan, K. C. (2019, October). A case study of predicting banking customer's behaviour by using data mining. In *2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC)* (pp. 1-6). IEEE.