



بسمه تعالی

بررسی کاربرد داده کاوی در جهت کاهش ریسک حاصل از تخصیص وام به مشتریان بانک

حامد هاشمی

چکیده

موسسه های اعتباری که بانک ها رکن اساسی آن شناخته می شوند، بخش بزرگی از در آمد خود را از طریق ارائه تسهیلات به مشتریان تامین می کنند. بدین منظور ابتدا با استفاده از سنجش پارامترهای مختلف، متقاضیان را اعتبارسنجی نموده و سپس با در نظر گرفتن ریسک اعتباری حاصل از انجام این کار، تصمیم های لازم در خصوص تخصیص یا عدم تخصیص اعتبار به متقاضی، میزان اعتبار قابل تخصیص، و همچنین نرخ سود متناسب گرفته خواهد شد.

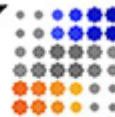
با توجه به اینکه در حال حاضر در کشور اعتبار سنجی مشتریان با استفاده از قضاوت نیروی انسانی صورت می پذیرد و راه های غیرانسانی پیشنهادی و حتی پیاده سازی شده در مقیاس پایلوت تا حال حاضر نتوانسته اند موفقیت چندانی پیدا کنند و همچنین نیاز به دانش صریح و صحیح از میان داده ها و اطلاعات موجود در جهت مدیریت ریسک و تخصیص آگاهانه ی وام به مشتریان جدید و گذشته به صورت روزافزون افزایش می یابد، از این رو راه رسیدن به یک راه حل غیرانسانی و محکم بر مبنای داده ها و اطلاعات موجود جهت اعتبارسنجی مشتریان بانک به منظور تخصیص وام هنوز به عنوان یک نیازمندی بدون پاسخ مناسب و بادوام در کشور باقی مانده است.

کلید واژه ها: داده کاوی، اعتبارسنجی مشتریان، الگوریتم های داده کاوی در اعتبارسنجی مشتریان، پارامترهای موثر اعتبارسنجی مشتریان،

مقدمه

با توجه به اینکه یکی از منابع اصلی درآمد بانک ها از طریق ارائه تسهیلات به مشتریان حقیقی و حقوقی تامین می گردد، مقوله وام دهی همواره یکی از نقاط مورد توجه فعالان این صنعت بوده است. اما همانطور که سود تخصیص تسهیلات توانسته است نقطه جذبی برای بانک ها به شمار رود، همواره نگرانی بازپس ندادن اصل و سود وام نیز در طول تاریخ، آن ها را تهدید کرده است. بنابراین بانک ها برای در اختیار قرار دادن انواع تسهیلات اعطایی به مشتریان خود، نیاز به انجام بررسی های کاملی به





به مشتریان بانک

منظور شناخت متقاضیان از ابعاد کمی و کیفی دارند تا از این طریق ارزیابی کاملی از سنجش توان بازپرداخت و محاسبه احتمال عدم بازپرداخت تسهیلات از سوی آنان را انجام دهند. از آنجائیکه این ارزیابی در حال حاضر در کشور صرفاً از طریق قضاوت نیروی انسانی صورت می پذیرد و احتمال خطا در آن بسیار بالا می باشد، بنابراین برنامه هایی که بتوانند تصمیم گیرندگان این بخش را در خصوص تصمیم گیری در امر تخصیص تسهیلات به هر مشتری یاری رسانند، می توانند به طور موثری نظرات مثبت آن ها را جلب نمایند.

با توجه به وجود منابع داده ای مناسبی در بانک که با ایجاد زیرساخت انباره داده و انباره های اختصاصی داده¹ برای داده ها و کسب و کارهای مختلف مانند تسهیلات، امکان پیاده سازی مناسبی از منظر امر داده کاوی فراهم گشته است، لذا می توان با استفاده از داده های خام و همچنین مقیاس های² ساخته شده در انباره داده تسهیلات، به دانش مفیدی در این خصوص دست پیدا نمود که با استفاده از این دانش می توان در جهت مدیریت ریسک گام های موثری برداشته و آگاهانه تر در خصوص تخصیص وام به مشتریان جدید و گذشته عمل کرده و ریسک عدم بازگشت اصل و سود تسهیلات بانکی را کاهش داد.

این تحقیق با موضوع ((بررسی کاربرد داده کاوی در جهت کاهش ریسک حاصل از تخصیص وام به مشتریان بانک)) سعی خواهد داشت سودمندی بالای استفاده از تکنولوژی داده کاوی در اعتبارسنجی مشتریان بانکی جهت تخصیص وام را نشان داده و با نشان دادن نتایج حاصل از برخی مقالات انتشار یافته که از انجام یک کار آزمایشگاهی بر روی یک بانک انجام شده است، اقدام به معرفی پارامترهای موثر در اعتبارسنجی متقاضیان تسهیلات بانکی نموده و برترین الگوریتم های به کارگرفته شده در این زمینه را مطابق با همان مطالعه های موردی مطرح نماییم.

مروری بر داده کاوی

امروزه حجم بالایی از داده ها در سازمان دولتی و موسسات علمی و تجاری در حال تجمیع و جمع آوری می باشد که به آن ها به عنوان یک سرمایه استراتژیک نگاه می شود. تکنولوژی های به روز مرتبط با جمع آوری اطلاعات و ذخیره سازی، این امکان را برای سازمان ها و موسسه ها فراهم کرده است که بتوانند داده ها و اطلاعات را جمع آوری و ذخیره نموده و از آن ها در برنامه های کاربردی در حوزه های مختلف دنیای کسب و کار، انجمن های علمی و انجمن های دارویی، بهره گیرند.

¹ Data Marts

² Measures



به مشتریان بانک

مجموعه فعالیت های درگیر در تحلیل این پایگاه های داده ای بزرگ، معمولاً با هدف استخراج اطلاعات مفید جهت پشتیبانی از تصمیم گیرندگان، به روش های مختلفی، از قبیل داده کاوی، اکتشاف دانش³، شناسایی الگوها⁴ و یادگیری ماشین⁵، انجام می شود.

به طور خاص واژه داده کاوی مشخص کننده ی فرآیند استخراج و تحلیل یک مجموعه داده، معمولاً با حجم زیاد، جهت شناسایی الگوهای با قاعده و معین، به منظور استخراج دانش مرتبط و به دست آوردن قوانین بازگشتی معنی دار می باشد.

داده کاوی در هر دو منظر مطالعات تئوری و برنامه های کاربردی، یک نقش همواره رو به رشد را بازی می کند. [1] در تعریفی دیگر آورده شده است که واژه داده کاوی نشان دهنده ی فرآیندی است که در خلال آن الگوهای کشف نشده موجود بین داده ها شناسایی می گردد. به واقع داده کاوی می تواند مسببی باشد برای اکتشاف دانش از پایگاه های داده ای موجود.

[2,5]

داده کاوی براساس روش های یادگیری استنتاجی بنا شده است، روش هایی که هدف اصلی آنها رسیدن به استنتاج قوانین عمومی با شروع از یک مجموعه ای از نمونه های در دسترس، که شامل مشاهدات گذشته ذخیره شده در یک یا تعداد بیشتری پایگاه داده است، می باشد. به عبارتی دیگر، هدف اصلی از تحلیل داده کاوی، ترسیم استنتاج های بیشتر بر روی نمونه مشاهدات گذشته و تعمیم این استنتاج ها به کل جامعه مورد مطالعه، با استفاده از روشی که موجب فراهم گشتن دقت هر چه بیشتر گردد، می باشد. مدل ها و الگوهای شناسایی شده در روش مورد استفاده برای استنتاج، می تواند به شیوه ها و شکل های مختلفی حاصل شود که در این زمینه می توان به معادله خطی، مجموعه قوانین قابل بیان به فرم "اگر- آنگاه- در غیر این صورت"⁶، خوشه بندی⁷، چارت ها و درخت ها اشاره نمود. [1]

مدل ها و روش های داده کاوی

روش های مختلف یادگیری برای انجام وظایف مختلف داده کاوی در دسترس می باشند. تعدادی از تکنیک های سرچشمه گرفته از فیلد علوم کامپیوتر، مانند درخت های دسته بندی⁸ یا قوانین وابستگی⁹، با عنوان یادگیری ماشین یا کشف دانش موجود در پایگاه های داده اشاره می شود. در اغلب موارد یک روش مبتنی بر تجربه، گرایش به غلبه کردن بر این گروه از تکنیک ها دارد.

³ Knowledge Discovery

⁴ Knowledge Discovery

⁵ Machine Learning

⁶ If-Then-Else

⁷ Clustering

⁸ Classification Trees

⁹ Association Rules

به مشتریان بانک

دیگر روش‌ها متعلق به آمارهای چندگانه، مانند رگرسیون یا دسته‌بندی کننده‌های بیزین^{۱۰}، می‌باشند که اغلب به ذات پارامتریک هستند اما بیشتر به صورت تئوری بنا شده‌اند. اما بیشتر توسعه‌های اخیر، دربرگیرنده‌ی روش‌های ریاضی یادگیری، مانند تئوری یادگیری آماری، که براساس پایه‌های تئوری محکم و قدرتمندی بنا شده‌اند و جایگاه خود را در محل برخورد اصول و قواعد مختلف که از تلاقی تئوری احتمال، تئوری بهینه‌سازی و آمار به دست می‌آید، بنا نهاده است. دسته‌بندی تکنیک‌های داده کاوی در منابع مختلف، کمی متفاوت می‌باشد. با توجه به مطالعات انجام شده، تمام منابع، بر

۴دسته زیر اتفاق نظر دارند:

- دسته بندی
- خوشه بندی
- وابستگی
- اکتشاف دنباله^{۱۱} [1,2,4]

تکنیک‌های داده کاوی دیگری نیز در منابع مختلف به ۴گروه فوق اضافه شده‌اند که در این خصوص می‌توان روش‌های زیر را

dataacademy.ir

نام برد:

- رگرسیون
- بصری سازی^{۱۲} [1,2,4]
- تکنیک پیش بینی^{۱۳} [2,4]
- تکنیک تفکیک و مشخص سازی^{۱۴} [1]

همچنین با توجه به دسته بندی که پیشتر خصوص انواع روش‌های یادگیری با ناظر و بدون ناظر عنوان شد، تکنیک‌های دسته بندی، اکتشاف دنباله، رگرسیون، تمایز و مشخص سازی را در گروه یادگیری با ناظر و تکنیک‌های خوشه بندی، بصری سازی و وابستگی را در گروه یادگیری بدون ناظر، قرار خواهند گرفت. [1]

¹⁰ Bayesian Classifiers

¹¹ Sequence Discovery

¹² Visualization

¹³ Forecasting

¹⁴ Characterization and Discrimination

به مشتریان بانک

اهمیت داده کاوی

در پاسخ به افزایش تقاضای مشتریان و رقابت سخت و دشوار موجود در بازار، سازمان ها می بایست از تکنولوژی های جدیدی استفاده نمایند که بتوانند حجم زیادی از داده ها را تولید و جمع آوری نمایند. حال دیدن مسائلی مانند ناسازگاری، تکراری بودن، حساسیت داده ها به زمان، وجود منابع داده ای مجزا و مختلف، و از بین رفتن بیش از حد داده ها، تنها بخش کوچکی از مشکلاتی است که اغلب تصمیم گیرندگان در سازمان ها را به خود مشغول ساخته است. در چنین جایی است که تکنولوژی های هوشمندسازی کسب و کار^{۱۵} که دربرگیرنده ی نرم افزارهایی است که می توانند انبوه داده ها را غربال کرده و به تکه های اطلاعاتی با بیشترین فایده برسند، به کمک تصمیم گیرندگان سازمانی می آید. یکی از تکنولوژی های هوشمندسازی کسب و کار، داده کاوی نام دارد که عمل آشکار سازی الگوها و ارتباطات مفید در بین داده های متنوع ذخیره شده را برعهده دارد. با گذشت زمان، سازمان ها و اقتصاد جهانی پیچیده تر شده و منابع عدم قطعیت نیز افزایش می یابد و داده کاوی نیز با استفاده از ترکیب علمی مانند آمار، کامپیوتر، یادگیری ماشین، هوش مصنوعی و دیگر تخصص های لازم، یک چارچوب برای مقابله با این عدم قطعیت ها، ارائه می دهد.

می بایست به این نکته توجه داشت که پاسخ دهی به سوالات ساده ای مانند "در فصل گذشته چقدر فروش داشته ام؟" و یا "فروش منطقه ی غربی در مقایسه با منطقه ی شرقی چه طور بوده است" که با نوشتن یک کوئری و یا با استفاده از نرم افزارهای گزارش ساز و یا یک مدل تحلیل های برپایه فرآیند تحلیل برخط به دست می آید، نیاز واقعی تصمیم گیرندگان سازمانی نمی باشد. این افراد با سوالات چالشی واقعی در دنیای کسب و کار روبرو می باشند. پاسخ سوالاتی مانند "چه مشتریانی احتمالاً سازمان را ترک می کنند؟، چرا؟ و چه وقت؟"، نیاز واقعی تصمیم گیرندگان می باشد که تکنیک های مدلسازی قوی موجود در داده کاوی، می تواند این نیاز را مرتفع سازد.

از سویی دیگر، دقت زیاد در مسائل تحلیلی برای شناسایی موثر نیازمندی های حجیم و رو به رشد کسب و کار، مانند تشخیص دقیق تر ریسک ها، افزایش بهینه ی قیمت و پیش بینی نیازمندی ها و اولویت های مشتری، مورد نیاز می باشد که این قبیل برنامه های کاربردی داده کاوی، به طور فزاینده ای باارزش بوده و احتمالاً در تمام سازمان ها خصوصاً سازمان های بزرگ، مورد نیاز می باشد.

در نهایت می توان به این موضوع اذعان داشت که در اقتصاد مشتق شده از داده که در دنیای کنونی وجود دارد، داده کاوی یک ابزار ضروری برای رسیدن به سودمندی بالا، کاهش عدم قطعیت، حفظ مشتریان، تخفیف ریسک، افزایش عایدی، پالایش فرآیندها

¹⁵ Business Intelligence



به مشتریان بانک

و بهینه سازی منابع تخصیص یافته محسوب می گردد و تمام این موارد نشان از اهمیت بالای جایگاه داده کاوی، در سازمان ها دارد.[5]

با توجه با آنچه که درخصوص اهمیت داده کاوی آورده شد، حال این سوال مطرح می شود که داده کاوی در چه حوزه هایی می تواند کاربرد داشته باشد. استفاده از این تکنولوژی بسیار وسیع بوده و نمونه های کاربردی آن را می توان در طیف گسترده ای از کسب و کارها از قبیل بانکداری، بیمه، پزشکی، پست، علوم نجوم و... مشاهده نمود.[2] متدولوژی های داده کاوی می توانند برای حوزه های گوناگون و متنوعی، از بازاریابی و کنترل فرآیند تولید تا مطالعه عوامل ریسک در تشخیص های درمانی، و از ارزیابی تاثیر داروهای جدید تا کشف تقلب^{۱۶} به کار گرفته شوند و تصمیم گیرندگان در هر حوزه را پشتیبانی نمایند. کاربردهای این متدولوژی ها بسیار فراگیر بوده و حتی می توان ردپای آن را در متن کاوی^{۱۷}، شناسایی تصویر^{۱۸}، و وب کاوی^{۱۹} نیز مشاهده نمود.[1]

داده کاوی در بانکداری

صنعت بانکداری به سبب ایجاد توانایی دریافت و ذخیره ی مجموعه انبوهی از داده ها و دسترس پذیری ابزارهای جدید برای تحلیل، به سرعت تغییر کرده است. ابزارهای جدید و متداولی که می توانید این پدیده تغییر را توصیف نمایند، دو ابزار انبار داده و داده کاوی می باشند.

تغییرات رخ داده در این صنعت، موجب گشته است که فعالیت های سنتی بانکداری مانند تخصیص تسهیلات در شعب بانک تغییر شکل یافته و حتی موجب متولد شدن برخی کسب و کارهای جدید در بانکداری گشته است. به عنوان مثال کسب و کار کارت اعتباری بدون استفاده از کامپیوترهای فوق العاده قدرتمند و متدهای آماری جدیدی که امروزه موجود می باشند، در گذشته غیرقابل ایجاد بوده اند.[5]

از آنجائیکه یکی از علت های مهم نیاز به داده کاوی در سازمان ها، به فشار بازار برمی گردد. در صنعت بانکداری نیز با توجه به وجود عنصر کلیدی مشتری و تاثیر بالای رقبا بر میزان سوددهی و همچنین جذب و حفظ مشتری، داده کاوی توانسته است تاثیر فراوانی بر روی این صنعت از خود قرار دهد.[2]

¹⁶ Fraud Detection

¹⁷ Text Mining

¹⁸ Image Recognition

¹⁹ Web Mining



به مشتریان بانک

مواردی مانند پیش بینی احتمال وفاداری مشتریان، درصد احتمال ریزش مشتریان، درصد احتمال استفاده از خدمات، پیش بینی مانده مشتریان در یک نوع سپرده سرمایه گذاری مشتریان، تاثیر بالای داده کاوی در مدیریت ارتباط با مشتریان²⁰، بخش بندی مشتریان و هدفمند سازی کمپین تبلیغات متناسب با هر سگمنت، پیش بینی احتمال وجود یا عدم وجود ریسک در بازگرداندن مبلغ تسهیلات توسط مشتری، درصد احتمال عدم بازگرداندن مبلغ تسهیلات توسط مشتری، شناسایی تقلب، وده ها نمونه ی دیگر، نشان دهنده ی کاربرد نرم افزارهای داده کاوی در این صنعت مهم در هر کشور می باشد.

با توجه به اینکه شرکت یا سازمانی که به طور موثر ابزارهای داده کاوی و تکنولوژی ها را به عنوان یک وسیله قدرت و نفوذ به کار می گیرد، می تواند به مزایای استراتژیک رقابتی دست پیدا کرده و آن ها را حفظ نماید، بنابراین یک بانک که یک موسسه اعتباری محسوب می شود، می تواند با استفاده از داده کاوی، موقعیت های جدیدی را به وسیله تبدیل داده ها به یک حربه استراتژیک، خلق کرده و از آن ها بهره برداری نمایند و این به معنای مزیت رقابتی یک بانک در بین سایر رقبا خواهد بود. اهمیت این موضوع به خصوص در بانک های خصوصی پررنگ تر خواهد بود.

مروری بر کاربردهای داده کاوی در بانک

اگرچه ما در این تحقیق صرفاً به بررسی داده کاوی در خصوص تسهیلات خواهیم پرداخت، اما نفوذ داده کاوی در تمام جنبه های صنعت بانکداری روز، غیرقابل انکار می باشد. بنابراین به برخی از برنامه های کاربردی مهم داده کاوی در این صنعت بزرگ، نگاهی خواهیم داشت.

• بازاریابی رابطه ای

در صنعت بانکداری که مشتریان سرمایه یک بانک را تشکیل می دهند و تعداد زیادی از مشتریان جدید از طریق تبلیغات و بازاریابی جذب می شوند، مفهوم بازاریابی برای یک بانک پررنگ خواهد بود.

برای تداوم حیات و حفظ مزیت رقابتی در بازارهایی با رقابت روزافزون و فزاینده، بسیاری از سازمان ها به سمت بازاریابی رابطه ای با تمرکز بر حداکثر نمودن ارزش دوره عمر مشتریان خود و مدیریت ریزش مشتریان روی آورده اند. در واقع تعداد بیشتری از سازمان ها متوجه شده اند که گرانبهاترین سرمایه آن ها، پایگاه مشتریان فعلی شان است. نگهداری مشتریان یک استراتژی ارزشمند است که سودآوری بلندمدت و موفقیت سازمان ها را تضمین می کند.

²⁰ Customer Relationship Management(CRM)



به مشتریان بانک

به طور کلی برنامه های کاربردی داده کاوی در فیلد بازاریابی رابطه ای به طور معنی داری در افزایش محبوبیت این متدولوژی ها شرکت داشته اند. برخی کاربردهای مرتبط در این بخش عبارت هستند از:

- شناسایی آن بخش از مشتریانی که احتمال بیشتری به پاسخگویی به کمپین های بازاریابی دارند.
- شناسایی بخش های مشتریان هدف برای حفظ و ابقاء کمپین ها
- پیشگویی نرخ پاسخ های مثبت به کمپین های بازاریابی
- درک و تفسیر رفتار خرید مشتریان
- تحلیل محصولاتی که با یکدیگر و همزمان توسط مشتریان خریداری می شوند. از اینگونه تحلیل ها با عنوان تحلیل سبد بازار یاد می گردد.

اما هدف از بازاریابی رابطه ای، افزایش مزایای رقابتی از طریق کسب رضایتمندی مشتریان است و این امر امکان پذیر نخواهد بود، مگر آنکه مدیریت ارتباط با مشتری در یک سازمان بنا نهاده شود. توسعه ی روابط با تک تک مشتریان می تواند در نهایت موجب افزایش وفاداری، نگهداری و همچنین سودآوری مشتری گردد. مدیریت ارتباط با مشتری دارای ۴ بعد می باشد:

- شناسایی مشتری dataacademy.ir
- جذب مشتری
- نگهداری مشتری
- توسعه ی مشتری

تمام این ۴ بعد هدف مشترکی دارند که همان ایجاد درک عمیق از مشتری، به منظور ماکزیموم کردن ارزش مشتری برای سازمان در دراز مدت است. این ۴ بعد که به عنوان چرخه ای بسته از سیستم مدیریت مشتری در نظر گرفته مس شود، می تواند تحت تاثیر تکنیک های داده کاوی قرار گرفته و داده کاوی به وسیله ی استخراج و یا تشخیص رفتارها و ویژگی های پنهان مشتریان از پایگاه داده های بزرگ، به انجام هدف مشترک آن ها کمک نماید. شکل زیر چارچوب طبقه بندی برای تکنیک های داده کاوی در مدیریت ارتباطات با مشتریان را نمایش می دهد.

در بخش شناسایی مشتری، آن دسته از مشتریانی که احتمال بیشتری دارد تا مشتری ما شوند و یا سود بیشتری را به همراه خواهند داشت، شناسایی می شوند. اجزای این بخش شامل تحلیل مشتری هدف و بخش بندی مشتری می باشد. پس از شناسایی سگمنت های مشتریان بالقوه، سازمان ها می توانند منابع و تلاش های خود را معطوف جذب این گروه از مشتریان نمایند.



به مشتریان بانک

بنابراین وارد بخش جذب مشتری می شویم. در این مرحله می توانیم از بازاریابی مستقیم بهره گیریم. در فاز بعد نگهداری از مشتری را خواهیم داشت و بعضی رضایتمندی مشتریان را پیش رو خواهیم داشت. اجزای این بخش شامل بازاریابی فرد به فرد، برنامه های وفاداری و مدیریت شکایات خواهد بود. در گام آخر نیز شاهد توسعه ی مشتری می باشد که دربرگیرنده ی تحلیل ارزش طول عمر مشتری، فروش پیوندی و تحلیل سبد بازار می باشد. هدف از این گام توسعه ی میزان تراکنش، ارزش تراکنش و سودرسانی هر مشتری می باشد. [6,7]

• کشف تقلب

کشف تقلب یکی دیگر از فیلهای مرتبط با کاربرد داده کاوی می باشد. تقلب ممکن است صنایع مختلف از قبیل صنعت بخش تلفن، بیمه (نمونه ی تقلب در این صنعت ادعاهای ناصحیح می باشد)، و بانکداری (استفاده غیرمجاز از کارت اعتباری و چک های بانکی و تراکنش های مالی غیرمجاز، نمونه هایی از تقلب در این صنعت می باشند) را تحت تاثیر قرار دهد.

یکی از مهمترین موانع برای استفاده از بانکداری اینترنتی عدم امنیت تراکنشها و برخی سوءاستفاده ها در مسیر انجام مبادلات مالی است. به همین دلیل جلوگیری از نفوذ غیرمجاز و تشخیص جرم از مسائل مهم در مؤسسات مالی و بانکهاست.

یکی از روشهای مورد استفاده در تشخیص جرم، استفاده از روشهای داده کاوی است که بر تحلیلهای آماری و کشف رفتار مشتریان و استفاده از الگوها برای شناسایی جرم تمرکز دارند. [8] این روشها مبتنی بر یادگیری قواعدی خاص هستند و قادرند شاخص های رفتارهای فریب آمیز را از پایگاه داده های بزرگ تراکنشهای کاربران کشف کنند. این شاخصها برای ایجاد سیستمهای پیشگیرنده²¹ استفاده می شوند تا رفتارهای غیرمعمول مشتریان را ثبت کرده و رفتارهای مشکوک را از میان آنها شناسایی کنند. در نهایت خروجی این سیستم ها می تواند برای اعلام هشدار و اخطار درخصوص کاربران متخلف استفاده شود.

قواعد وابستگی نیز یکی از بهترین روشهای داده کاوی برای خلق چنین مدلهایی است. از این روش برای استخراج دانش و به دست آوردن الگوهای رفتاری غیرمتعارف از مجموعه بزرگ تراکنشهای کاربران در پایگاه داده های تراکنش کارت اعتباری برای تشخیص و جلوگیری از وقوع جرم استفاده شده است. روش دیگری که تاکنون برای شناسایی و تشخیص جرم استفاده شده شبکه های عصبی مصنوعی است که قابلیت استخراج الگو از پایگاه داده های حاوی تراکنشهای گذشته مشتریان را دارند. این شبکه ها آموزش پذیر بوده و قابلیت انطباق با شکلهای جدید جرم را دارا هستند. [9]

²¹ Monitoring

به مشتریان بانک

• امتیازدهی اعتباری^{۲۲}

در اغلب فرم های ساده، امتیازدهی اعتباری می تواند به عنوان تکنیکی کمک رسان به کسانی که در خصوص اعطای اعتبار تصمیم گیرنده می باشند، مطرح گردد.

مدل های امتیازدهی اعتباری مانند همه ی ابزارهای پیش بینی طراحی شده برپایه ی آمار، براساس این فلسفه ی ساده می باشد: "گذشته، آینده را پیش بینی می کند". اینگونه مدل های آماری، رفتار گذشته و شرایط فعلی را تحلیل کرده و ریسک آتی را پیش بینی می نمایند. به طور معمول یک مدل مبتنی بر امتیازدهی اعتباری بدین منظور طراحی شده است که احتمال بد اعتباری یک مشتری و یا احتمال اینکه از نوع مشتریانی باشد که پرداخت های خود را به موقع انجام نمی دهد را پیش بینی نماید، بنابراین مشخص است که با استفاده از این مدل، تصمیم گیرندگان می توانند با در نظر گرفتن ریسک موجود، بهتر و سریعتر در خصوص اینگونه مسائل مالی، تصمیم گیری نمایند.[4]

البته اگرچه می توان از مدل امتیازدهی اعتباری برای موارد بسیاری استفاده نمود، اما به جز معمولترین مورد استفاده ی این مدل که به مساله اعطا یا عدم اعطای وام در موسسات مالی باز می گردد، بیشترین استفاده از آن در تصمیم گیری جهت تهیه مسکن، بیمه و حتی استخدام کارمند بوده است. به واقع با این مدل مطابق با کسب و کار پیش رو، به افراد امتیازی تعلق گرفته و بر مبنای آن، تصمیم نهایی گرفته می شود.

استفاده از امتیازدهی اعتباری برای تشخیص میزان تعهد یک مشتری، سال ها است که در امر وام دهی به مشتری مورد استفاده قرار می گیرد، اما این مدل حتی محدود به تصمیم گیری جهت رسیدن به پاسخ مثبت یا منفی در امر تخصیص وام نبوده، و امروزه در بسیاری از بانک ها، در خصوص محدودیت مبلغ قابل اعطا به هر مشتری نیز از این مدل استفاده می نمایند. امروزه تکنیک های فراوانی برای ساخت مدل های اعتباری مورد استفاده قرار می گیرد که معمولترین آن ها، از نوع تکنیک های آماری سنتی می باشند. به عنوان مثال، برخی از مدل های اولیه امتیازدهی اعتباری با استفاده از تحلیل تفکیکی^{۲۳}، توسعه یافته اند. همچنین می توان به الگوریتم ژنتیک، رگرسیون، برنامه نویسی خطی، و سیستم های خبره نیز که برای توسعه ی مدل امتیازدهی اعتباری مورد استفاده قرار می گیرند نیز اشاره نمود، اما در سال های اخیر، تکنیک های داده کاوی به طور فزاینده ای جهت ساخت مدل های امتیازدهی اعتباری، مورد استفاده قرار گرفته اند. به خصوص الگوریتم درخت تصمیم به علت قابلیت تفسیر و نمایش راحت اطلاعات، به عنوان یک تکنیک محبوب در این امر شناخته شده است.

²² Credit Scoring

²³ Discriminate Analysis



به مشتریان بانک

بهرحال، علیرغم وجود مقاومت اولیه مشاهده شده توسط وام دهندگان تجاری در خصوص استفاده از امتیازدهی اعتباری، رشد موجود در بازار وام دهی، یک محیط کامل را برای قدرت نفوذ کاربردهای این مدل، ایجاد کرده است. [3, 6]

• ارزیابی ریسک

هدف از تحلیل ریسک، ارزیابی ریسک های مرتبط با تصمیم های آتی می باشد. به عنوان مثال، یک بانک ممکن است با استفاده از مشاهدات گذشته و در دسترس، مدل پیشگویی را متناسب با تصمیم گیری در خصوص اعطا یا عدم تخصیص تسهیلات براساس خصوصیات متقاضی، توسعه دهد.

• سایر کاربردها

از دیگر کاربرد های داده کاوی در صنعت بانکداری، می توان به کشف پول شویی و تشخیص مشتریان ثابت و همیشگی اشاره کرد.

تأثیر داده کاوی در اعتبارسنجی مشتریان بانکی جهت تخصیص وام

با توجه به آنچه که از اهمیت تکنولوژی داده کاوی و نفوذ آن در صنعت بانکداری گفته شد، تاثیر بالای این تکنولوژی در این خصوص امری انکارناپذیر است. اعتبارسنجی متقاضیان تسهیلات به عنوان یکی از حوزه های سود آور و یا زیان آور در بانک ها نیز می تواند به عنوان یک مبحث قابل کار مورد بررسی قرار گیرد.

در صورتیکه بتوان یک مدل داده کاوی کارآمدی در این زمینه ارائه کرد، ادعای برداشت گامی موثر به سوی بهبود ساخت تصمیم به حق می باشد. اما نکته ی طلایی در این است که تصمیم نه تنها از جهت بررسی ریسک حاصل از تخصیص تسهیلات، بلکه از جهت مقدار پولی قابل پرداخت به متقاضی با رعایت میزان ریسک قابل قبول برای سازمان نیز می تواند مطرح شود. علاوه بر آن قابلیت شفاف سازی دلایل عدم امکان ارائه ی تسهیلات و یا تخصیص با مبلغ محدود به خود شخص متقاضی و سازمان های ذیصلاح و مرتبط و همچنین خلق موقعیت های جدید به وسیله تبدیل داده ها به یک حربه استراتژیک و بهره برداری از مزیت رقابتی در میان سایر رقبا نیز از مزایای حاصل از به کارگیری داده کاوی در اعتبارسنجی متقاضیان محسوب می گردد.

داده کاوی و مقوله ی تخصیص وام بانکی در دنیا

با توجه به اینکه این مقاله ی تحقیقاتی به بررسی کاربرد داده کاوی در مقوله ی تخصیص وام می پردازد، بهتر است نگاهی داشته باشیم بر مطالعات انجام شده درباره ی این موضوع در بانک های ایران و جهان. اما غالب مقاله های منتشر شده و در دسترس که به مطالعه ی موردی پرداخته اند، موضوعات مورد بحث خود را در دو گروه دنبال می کردند، نخست محاسبه ی امتیازدهی

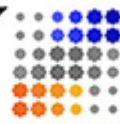


به مشتریان بانک

اعتباری که پیشتر در مورد آن توضیح داده شد و دوم به ساخت مدل های پیشگویی گوناگون با استفاده از تکنیک های مختلف داده کاوی پرداخته و در نهایت با مقایسه ای که انجام می دهند، یکی از مدل ها را به عنوان مدلی با دقت بیشتر پیشنهاد می دهند.

از این رو ما در این بخش به بررسی دو مقاله مربوط به دو کشور مختلف پرداخته ایم که نتیجه ی بررسی هر کدام را در ستون آخر بیان کرده و سعی کرده ایم نتایج مختلفی که هریک از محققان از انجام کار خود برآورد کرده اند را نمایش دهیم. بنابراین با نگاهی بر خلاصه ی این مقاله ها می توانیم ببینیم محققان این حوزه با چه نوع دیدی به به مسئله ی اعتبارسنجی متقاضیان تسهیلات بانکی پرداخته اند و چه خصیصه هایی را به عنوان پارامترهای موثر در بررسی اعتبار مشتریان مطرح کرده و از چه تکنیک ها و الگوریتم های داده کاوی در این امر کمک گرفته و در نهایت نتایج حاصل از انجام کار آزمایشگاهی آن ها چه بوده است. حاصل این تحقیق می تواند به طور موثری برای مطالعه های آتی با استفاده از تکنیک های داده کاوی مورد استفاده قرار گیرد.

| بانک مورد بررسی | نوع مساله | تکنیک داده کاوی مورد استفاده | الگوریتم مورد استفاده | پارامترهای مجموعه داده ی ورودی | نتیجه ی بررسی |
|--------------------------------|---|------------------------------|-----------------------|---|--|
| یک بانک تجاری از جنوب کشور چین | بررسی اعتبار مشتریان جهت تخصیص وام مسکن | دسته بندی | درخت تصمیم گیری | سن، جنسیت، میزان تحصیلات، شغل، وضعیت تاهل، میزان سابقه ی کاری، در آمد، وضعیت منزل مسکونی از نقطه نظر صاحبخانه یا موجر بودن، وضعیت سکونت از منظر ساکن بودن در یک مکان برای چندین سال | موثرترین پارامترهای که در در سه سطح اول ظاهر شدند، عبارت هستند از: وضعیت سکونت از منظر ساکن بودن در یک مکان برای چندین سال، درآمد و میزان تحصیلات [10] |



به مشتریان بانک

| | | | | | |
|---|---|---|--|---|-----------------------------|
| نتایج بررسی این تحقیق نشان می دهد که شبکه ی عصبی سه لایه با استراتژی سریع، از بین سایر تکنیک ها و الگوریتم های مورد بررسی، بیشترین میزان دقت را داشته است. [11] | سن، تاریخ ورود به بانک، تحصیلات، جنسیت، وضعیت تاهل، نوع تسهیلات، روش بازپرداخت تسهیلات، میزان سود وام، مدت زمان قرارداد، میزان تسهیلات اعطایی، نوع وثیقه و میزان مبلغ وثیقه | درخت تصمیم گیری با استفاده از الگوریتم چاید ^{۲۴} ، رگرسیون منطقی چند جمله ای و شبکه های عصبی با استراتژی سریع ^{۲۵} ، پویا ^{۲۶} و چندگانه ^{۲۷} | دسته بندی، شبکه های عصبی، درخت تصمیم و رگرسیون | به دست آوردن بهترین مدل پیشگو در محاسبه ی ریسک اعتباری مشتریان در بازپرداخت به بانک | یک بانک خصوصی از کشور ایران |
|---|---|---|--|---|-----------------------------|

جدول ۱- نتایج بررسی مقاله های انجام شده در زمینه ی ارتباط داده کاوی و تسهیلات

dataacademy.ir

نتیجه گیری

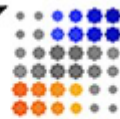
با توجه به اینکه یکی از منابع اصلی درآمد بانک ها از طریق ارائه تسهیلات به مشتریان حقیقی و حقوقی تامین می گردد، مقوله وام دهی همواره یکی از نقاط مورد توجه فعالان این صنعت بوده است. اما در مقابل فرصت درآمدزایی مقوله وام دهی برای این موسسه های اعتباری، همواره تهدید بالقوه بازپس ندادن اصل و سود وام توسط گیرنده ی تسهیلات، آن ها را تهدید کرده است. بنابراین جهت کاهش ریسک حاصل از تخصیص وام برای موسسه های اعتباری و بانک ها، نیاز به اعتبارسنجی متقاضیان تسهیلات از ابعاد کمی و کیفی ملموس و مشهود می باشد. در حال حاضر در کشور از قضاوت نیروی انسانی به عنوان راه حل بهره می برند، بنابراین اگر بتوانیم سیستم پشتیبانی از تصمیم که یک سیستم کامپیوتری می باشد را ارائه نماییم، توانسته ایم نظر مثبت تصمیم گیرندگان را تامین نماییم و این افراد می توانند نه تنها به اتکال قضاوت نیروی انسانی تصمیم خود را نهائی سازند، بلکه می توانند از هشدار عواقب اعطای چنین تسهیلاتی، به روش مکانیزه بهره مند شوند.

²⁴ CHAID

²⁵ Quick Strategy

²⁶ Dynamic Strategy

²⁷ Multiple Strategy



به مشتریان بانک

با توجه به آنکه تکنولوژی داده کاوی می تواند اطلاعات مفید از پایگاه های داده ای بزرگ را استخراج نماید، بنابراین گزینه ی خوبی برای پیاده سازی سیستم پشتیبانی از تصمیم می باشد. همچنین با توجه به وجود انواع روش های داده کاوی و الگوریتم های پیاده سازی و سرعت بالای به روز شدن آن، محققان و تحلیل گران از آن به عنوان برگ برنده ی عصر فن آوری اطلاعات کنونی یاد کرده و از آن در انواع حوزه های کاربردی کسب و کار و دنیای تجارت، بهداشت و علم بهره می برند. صنعت گران بانکی نیز از این مقوله مستثنی نبوده و سعی در استفاده از آن در انواع کاربردهای بانکی مانند مقوله ی اعتبار سنجی مشتریان جهت اعطای وام، بخش بندی مشتریان، کشف تقلب و بسیاری موارد دیگر دارند که در این تحقیق نیز سعی بر آن شد اهمیت و کاربرد این تکنولوژی روز دنیا در امر تخصیص وام بانکی و پیش بینی ریسک بازگشت یا عدم بازگشت اصل و سود وام را بررسی نموده و آنچه در تحقیق های گذشته که بر روی یک مطالعه ی موردی بنا شده اند، به عنوان پارامترهای موثر و همچنین الگوریتم های مناسب به کار گرفته شده در مدل های داده کاوی جهت رتبه بندی متقاضیان تسهیلات بانکی معرفی شده بودند را بررسی نماییم. همانطور که دیده شد، از هر دو نوع متغیر مالی و دموگرافیک بهره گرفته شده و درخت تصمیم از محبوبیت بالایی برخوردار بود.

dataacademy.ir