

باسمه تعالی

طرح پیشنهادی پایان نامه
(کارشناسی ارشد و دکتری)

رشته تحصیلی:

مقطع تحصیلی:

نوع دوره:

عنوان پایان نامه

فارسی: پیش بینی میزان ریسک شرکت های بیمه گذار بیمه درمان با رویکرد داده کاوی

انگلیسی:

Developing a Model Using Data Mining in order to Predicting Risk Amount of Health Insurance's Policyholder

۱. بررسی مطالعات پیشین (پژوهش‌ها . پایان‌نامه‌ها)

سال انجام پژوهش	پژوهشگر	عنوان پژوهش	نتایج
۲۰۱۳	Lookman Sithic	مروری بر تکنیک‌های داده‌کاوی استفاده شده برای کشف تقلب در بیمه	این تحقیق جزئیات بیشتری را از تقلب‌ها در زمینه بیمه و راه‌های آن با استفاده از روش‌های داده‌کاوی را مرور می‌کند.
۲۰۰۷	R.J. Kuo	بررسی قواعد انجمنی همراه با تحلیل خوشه و سیستم کلونی مورچگان برای پایگاه داده بیمه درمان در تایوان	نتایج نشان می‌دهد که روش پیشنهادی به تنهایی نمی‌تواند قوانین را با سرعت بیشتری از پایگاه داده استخراج کند بلکه موارد کلیدی و مهمتری را کشف می‌کند.
۲۰۰۲	Inna Kolyshkina	رویکردهای داده‌کاوی به منظور مدل کردن ریسک در بیمه	سن افراد تنها متغیری نیست که میزان ریسک به آن وابسته است.
۲۰۰۱	Young Moon Chae	رویکرد داده‌کاوی در تحلیل سیاست در حوزه بیمه درمان	الگوریتم‌های درخت تصمیم استفاده شده در این حوزه عملکرد بهتری نسبت به رگرسیون ارائه می‌دهند.

۲. بیان مسأله :

در گذشته، سازمان‌ها بدون داشتن هیچ‌گونه اطلاعاتی درباره مشتریان خود، به فروش محصولات و خدمات می‌پرداختند. امروزه با افزایش رقابت و تعداد رقبای بازار، نقش مشتریان در سازمان‌ها تغییر یافته است. مشتریان در تولید کالا و ارائه خدمات، فرایندها و توسعه دانش و توان رقابتی همراه و همگام اعضای سازمان هستند. بنابراین مدیریت مؤثر ارتباط با مشتریان جزئی حیاتی از سازمان‌ها به شمار می‌رود. همچنین سازمان‌ها درک کرده‌اند که حفظ مشتریان فعلی به دلایل مختلفی بسیار پرسودتر از جذب مداوم مشتریان جدید است. یکی از این دلایل آن است که جذب مشتری جدید بسیار پرهزینه‌تر و مشکل‌تر از حفظ

مشتریان فعلی است. به طوری که گفته شده است فروش محصول به یک مشتری جدید حدود پنج برابر، فروش محصول به یک مشتری فعلی هزینه بر است. از طرف دیگر، حفظ مشتریان فعلی باعث افزایش سود سازمان خواهد شد زیرا مشتریانی که مدت طولانی با سازمان در ارتباط هستند، در مقابل تبلیغات رقبا مقاوم تر می‌باشند و همچنین از طریق انتشار سخن مثبت^۱ باعث تبلیغ برای سازمان و جذب مشتریان بیشتر خواهند شد (Xie, Li, Ngai, & Ying, 2009).

از طرفی در دهه اخیر، با رشد فناوری اطلاعات و بروز ابزارهای جدید، سازمان‌ها اطلاعات بسیار زیادی از مشتریان در پایگاه‌های داده خود ذخیره نموده‌اند. تحلیل این داده‌ها و استخراج الگوهای مناسب از آن، به تصمیم گیرندگان سازمان کمک می‌نماید تا شناخت بهتری از مشتریان به دست آورده و بتوانند تصمیمات بهتری برای مدیریت ارتباط با آنان بگیرند. داده کاوی^۲ فرایند استخراج دانش مورد نیاز از این داده‌هاست که برای کشف اطلاعات در رابطه با مشتریان کاربردهای فراوانی دارد (Rygielski, Wang, & Yen, 2002).

همانند سایر بخش‌های اقتصادی، صنعت بیمه نیز شاهد تغییرات بسیاری در عرصه فناوری اطلاعات در طی سالهای گذشته بوده است. پیشرفت در زمینه‌های سخت افزاری و نرم افزاری فواید بسیاری را از قبیل کاهش هزینه‌ها، کاهش زمان پردازش داده‌ها، افزایش قابلیت‌های سودآوری، همزمان با رویارویی با چالش‌های جدید در فضایی که هر روز شاهد افزایش میزان رقابت در آن بوده‌ایم، به همراه داشته است. ابداعات و نوآوری‌های فنی از قبیل استخراج و نیز نگهداری داده‌ها به طور قابل ملاحظه‌ای موجبات کاهش هزینه‌های نگهداری، دستیابی و نیز پردازش داده‌ها را فراهم آورده است، بسیاری از سؤالات تجاری که در گذشته به علت کمبود داده‌ها و امکانات پردازشگری غیر ممکن، غیر عملی و غیر قابل حل می‌نمودند، در حال حاضر با استفاده از تکنیک‌های داده کاوی قابل بررسی و پاسخگویی هستند. یکی از شاخه‌های مهم محصولات شرکت‌های بیمه‌ای، بیمه‌نامه‌های اشخاص می‌باشد که مستقیماً زندگی بیمه‌گزاران را تحت پوشش قرار می‌دهد و از میان بیمه‌نامه‌های اشخاص بیمه‌نامه‌های درمان^۳ به دلیل تنوع پوشش‌ها، قراردادهای و ضریب خسارت بالایی که در شرکت‌های بیمه‌ای دارد از اهمیت بسیار زیادی برخوردار می‌باشد (Y M Chae, 2001).

¹ Word-of-mouth

² Data Mining

³ Health Insurance

- بیمه

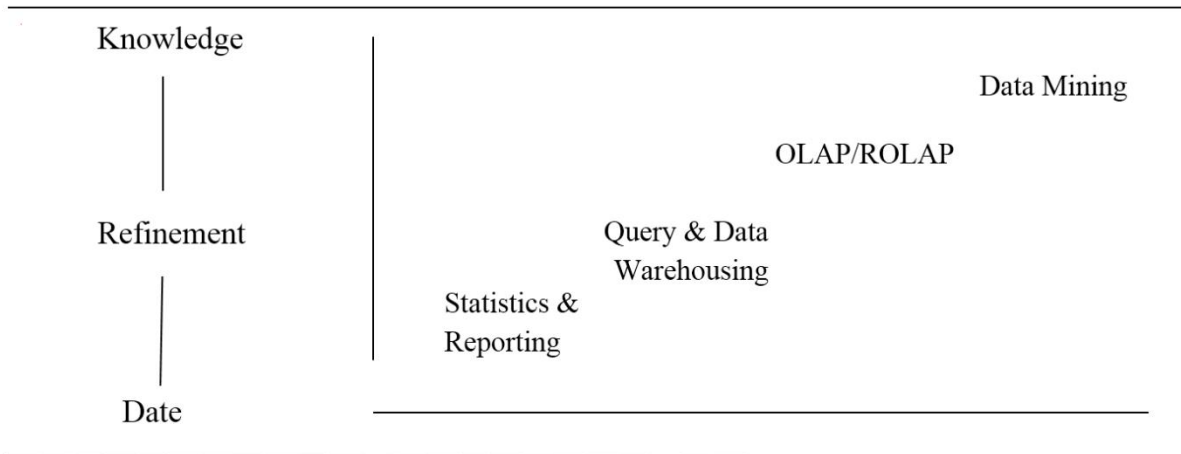
بیمه در حقیقت فرآیند یک واسطه‌گری مالی با ماهیت تعاونی می‌باشد که در آن چرخه تولید معکوس شده و مصرف‌کننده قبل از آنکه خدماتی دریافت کند بهای آن را می‌پردازد اما تعریف حقوقی بیمه طبق ماده یک قانون بیمه عبارت است از:

بیمه عقدی است که به موجب آن یک طرف تعهد می‌کند در ازای پرداخت وجه یا وجوهی از طرف دیگر در صورت وقوع یا بروز حادثه خسارت وارد بر او را جبران نموده و یا وجه معینی را بپردازد متعهد را بیمه‌گر و طرف تعهد را بیمه‌گذار، وجهی را که بیمه‌گذار به بیمه‌گر می‌پردازد حق بیمه و آنچه را که بیمه می‌شود موضوع بیمه گویند. به طور کلی تعاریف متعددی از بیمه ارائه شده است ولی تعریفی که از بیمه در رابطه با ریسک ارائه شده است عبارت است از بیمه سازو کار و یا خدمتی است برای انتقال ریسکهای مشخص بیمه‌گذار به بیمه‌گر در قبال پرداخت حق بیمه توافق شده.

بیمه در ساده‌ترین تعریف روشی است برای انتقال ریسک. در تعریف ماده یک قانون بیمه ایران، بیمه عبارت است از قراردادی که به موجب آن یک طرف (بیمه‌گر) تعهد می‌کند در ازای پرداخت وجه یا وجوهی از طرف دیگر (بیمه‌گذار) در صورت وقوع یا بروز حادثه خسارت وارده بر او را جبران نموده و یا وجه معینی را بپردازد. متعهد را بیمه‌گر، طرف تعهد را بیمه‌گذار و وجهی را که بیمه‌گذار به بیمه‌گر می‌پردازد حق بیمه و آنچه را که بیمه می‌شود موضوع بیمه نامند.

- داده کاوی

امروزه کمبود داده برخلاف گذشته مشکل کسب و کارهای کنونی نیست، بلکه مسئله حیاتی توانایی استخراج اطلاعات مفید است. به دلیل افزایش ورود اطلاعات به سبب فراهم بودن بسترهای ارتباطی و نیز وجود رایانه‌های پرسرعت و لزوم پردازش آنها توسط مدیریت و اتخاذ تصمیمات به موقع، استفاده از روش‌های آماری به عنوان ابزاری جهت استخراج اطلاعات مفید از حجم انبوهی از داده، افزایش یافته است. روش‌های آماری پدیده‌های غیر ساخت یافته و غیر سازماندهی شده را در قالبی ساخت یافته و سازماندهی شده نشان می‌دهند. شکل ۱ مراحل پالایش داده را نشان می‌دهد (Lejeune, 2001).



شکل ۱: روند تکامل جستجو برای کسب اطلاعات (Lejeune, 2001)

در واقع رشد سریع در دسترسی داده و تکنولوژی‌های پایگاه داده نیاز به ایجاد روش‌ها و ابزارهای جدید تبدیل داده به اطلاعات و دانش مفید به صورتی هوشمندانه و اتوماتیک را دو چندان نموده است. بنابراین داده کاوی به عنوان حوزه‌ای تحقیقاتی از اهمیت زیادی برخوردار است (Lee & Siau, 2001). روش‌های داده کاوی نتیجه سال‌ها تحقیق و توسعه محصولات است و منشأ پیدایش آن نیز به اولین سال‌های ذخیره سازی داده در رایانه‌ها بر می‌گردد (Rygielski et al., 2002).

- داده کاوی در صنعت بیمه

در صنعت بیمه داده کاوی به شرکت‌های بیمه جهت کسب سود بازرگانی بیشتر کمک می‌کند. برای مثال با استفاده از تکنیک استخراج داده‌ها، شرکت‌های بیمه می‌توانند با بهره‌برداری کامل از الگوهای رفتاری مشتریان خود و با شناخت بیشتری که از انگیزه‌های آنها به دست می‌آورند به کاهش میزان تقلب، افزایش امکان پیش‌بینی منابع تقاضا، افزایش فراگیری و نیز جلوگیری از کاهش مشتریان کمک نمایند (Y M Chae, 2001).

روش داده کاوی اغلب می‌تواند امکان اصلاح دستاوردهای آماری سنتی را در جهت حل مسائل بازرگانی فراهم آورد. برای مثال استفاده از رگرسیون خطی امکان حل مسائل بازرگانی را در راستای نیاز بیمه‌گران و به همراه مدل‌ها و پارامترهای قابل تفسیر مهیا می‌نماید. امکان اصلاح مدل‌های حاضر در اغلب موارد با استفاده از داده‌ها و از طریق کشف متغیرهای جدید، شناخت روابط و نیز کشف روابط غیرخطی فراهم می‌شود. مدل‌هایی که قادر به تخمین و پیش‌بینی روابط و رفتارها باشند می‌توانند سود بیشتری را با هزینه‌های کمتر حاصل نمایند.

۲. تشریح و بیان موضوع:

با تمرکز بر این توضیحات این تحقیق در نظر دارد با کاربرد عملی داده کاوی در یک شرکت بیمه بر روی مشتریان بیمه‌نامه‌های درمان بررسی نماید که مشتریان چگونه با تکنیک‌های داده کاوی در طبقه‌بندی مشتریان، قرار می‌گیرند. آیا از این طریق می‌تواند به شرکت‌های بیمه کمک کند که با شناخت دسته‌بندی‌های مختلف مشتریان و ویژگی‌های آنها از یک طرف و دسته‌بندی خدمات شرکت بیمه و ویژگی‌های آن از طرف دیگر بتوانند الگوی مناسبی جهت ارائه خدمات شخصی‌سازی شده مناسب به هر مشتری بر اساس ویژگی‌های منحصر به فرد او تبیین نمایند و خدمات مناسب را جهت کسب حداکثر سود سازمان و رضایت مشتری ارائه دهند و به این ترتیب حداکثر ارزش افزوده را از برقراری رابطه با مشتری کسب نمایند.

۴. ضرورت انجام پژوهش:

گسترش روزافزون فناوری اطلاعات و رقابتی شدن بازار، تولید محصولات باکیفیت به صورت انبوه، برای بقا و موفقیت سازمان‌ها کافی نیست و با افزایش اطلاعات مشتریان، فرآیند بازاریابی هر روز پیچیده‌تر و حساس‌تر می‌شود. همچنین، توجه به مبحث مدیریت ارتباط با مشتری هر روز اهمیت بیشتری پیدا می‌کند و بودجه اختصاص یافته به آن در سازمان‌ها افزایش می‌یابد.

با توجه به این موضوعات امروزه سازمان‌ها با ایجاد پایگاه‌های اطلاعاتی سعی به هنگام نمودن اطلاعات خود از مشتریان کرده و از این ابزار برای به دست آوردن مزیت رقابتی استفاده می‌کنند. در واقع سازمان‌ها با استفاده از ابزارهای داده‌کاوی در پی به دست آوردن الگوهای پنهان رفتار مشتریان در پایگاه‌های داده خود می‌باشند تا بتوانند با استفاده از این الگوها برای تصمیم‌گیری‌های مؤثر در زمینه ارتباط با مشتریان پردازند.

صنعت بیمه، یکی از صنایع کلیدی و پرقابلیت در دنیای امروزی می‌باشد، اهمیت این صنعت تا حدی است که به عنوان شاخصی در بررسی توسعه یافتگی کشورها محاسبه می‌شود. در کشور ایران ورود بیمه‌نامه‌های جدید با پوشش‌های متنوع و گسترده در بازار رقابتی امروز صنعت بیمه را به چالش کشیده است. با توجه به اینکه شرکت‌های بیمه‌ای، شرکت‌هایی خدماتی هستند، سرمایه اصلی این شرکت‌ها مشتریان می‌باشند. ورود شرکت‌های خصوصی از سال ۸۱ به بازار بیمه و به خصوص ورود دو شرکت دولتی (بیمه البرز و بیمه آسیا) به بخش خصوصی در سال ۸۸ رقابت بین شرکت‌ها را افزایش داده است. بنابراین توجه به الگوهای رفتاری مشتریان برای شرکت‌های بیمه امری ضروری تلقی می‌گردد. به همین دلیل سازمان‌ها نیاز دارند مشتریان خود را بشناسند و برای آنان برنامه‌ریزی کنند تا بتوانند با تصمیم راهبردی مناسب و اعمال صحیح آن موفقیت تجاری خود را بهبود بخشند. از طرفی با توجه به پیشرفت سریع فناوری اطلاعات، حجم اطلاعات ذخیره شده در پایگاه‌های داده شرکت‌های بیمه به سرعت در حال افزایش بوده و این پایگاه‌های داده بزرگ، حاوی حجم زیادی از داده‌ها و فرصت‌های قابل استفاده و بالقوه از اطلاعات تجاری با ارزش هستند. حجم بسیار زیادی از این داده‌ها مربوط به بیمه شده‌های درمان می‌باشد و از طرفی به دلیل رقابت فراوان در زمینه فروش بیمه درمان بین شرکت‌های بیمه‌ای و همچنین بالا رفتن روز افزون تقاضا برای این نوع بیمه‌نامه، حفظ مشتریان این نوع محصول و جلوگیری از چرخش آنان به سوی رقبا از مسائل مهم در شرکت‌های بیمه‌ای می‌باشد.

بنابراین این تحقیق به دنبال یافتن اطلاعات ارزشمند پنهان پایگاه داده‌های شرکت‌های بیمه‌ای و دسته‌بندی مشتریان بیمه‌نامه‌های درمان می‌باشد. تاکنون از برخی روش‌های آماری و یادگیری ماشینی برای این منظور استفاده شده است که البته این روش‌ها به تنهایی دارای محدودیت‌هایی هستند که در این تحقیق سعی

شده است تا با بهره‌گیری از روش‌های مختلف داده‌کاوی تا حد ممکن این محدودیت‌ها را از بین برده و برطبق آن، چارچوبی برای دسته‌بندی مشتریان بیمه ارائه شود. درواقع، هدف این است تا مشتریان بیمه نامه‌های درمان را که بیشتر به یکدیگر شبیه هستند دسته‌بندی کرده تا بتوان برای هر یک از دسته‌ها با توجه به ویژگی‌ها و ضریب خسارت هر دسته، قرارداد مناسب بیمه‌نامه درمان با پوشش‌های مناسب طراحی شود. بنابراین مشتریان جدید بر حسب ویژگی‌هایشان میزان خسارتشان پیش‌بینی می‌شود و در دسته مناسب قرار می‌گیرند و قرارداد مناسب بیمه‌ای برای آنها به کار گرفته می‌شود.

۵. اهداف پژوهش:

- تعیین و شناسایی شاخص ها در پیش بینی میزان ریسک بیمه گذاران
- ایجاد مدلی بر مبنای داده کاوی به منظور پیش بینی هر چه دقیق تر میزان ریسک بیمه گذاران

۶. سئوالات یا فرضیه‌های پژوهش:

- مدل پیش بینی میزان ریسک بیمه گذار بر مبنای داده کاوی چیست؟
- شاخص ها در پیش بینی میزان ریسک بیمه گذاران کدام است؟

۷. روش پژوهش :

۷-۱. روش‌شناسی پژوهش و دلیل به‌کارگیری آن

از آنجا که پایگاه داده مشتریان بیمه‌نامه درمان شرکت بیمه شامل مشخصه‌های جمعیت شناختی مشتری و مشخصه‌های مالی آن می‌باشد و تعداد این مشخصه‌ها بسیار زیاد است. بنابراین ابتدا براساس مصاحبه و جلسات با مدیران فنی و بررسی پژوهش‌های مشابه قبلی مهمترین مشخصه‌های جمعیت شناختی و مالی مشتری که جهت دسته‌بندی مشتریان و یافتن روابطی پنهان در میان داده‌ها برای پیش‌بینی‌های آتی، در این پژوهش مناسب‌ترین انتخاب می‌گردد. پیاده‌سازی مدل داده‌کاوی با استفاده از داده‌های آماده شده و الگوریتم‌های مورد نظر انجام می‌گیرد.

۷-۲. قلمرو پژوهش (موضوعی، مکانی و زمانی)

پژوهش حاضر مطالعه‌ای در حوزه‌ی کارگزاران بیمه در ایران انجام می‌دهد.

۷-۳. جامعه آماری و دلیل انتخاب آن

جامعه آماری مورد استفاده در این پژوهش، مجموعه داده‌های واقعی مربوط به مشتریان شرکت بیمه می‌باشد. این داده‌ها شامل فیلدهای اطلاعاتی مربوط به ویژگی‌های مشتریان و تراکنش‌های مالی مشتریان شامل حق بیمه، خسارت، پوشش‌های بیمه‌ای و ... می‌باشد.

۷-۴. نمونه آماری تعداد (نمونه)، روش نمونه‌گیری و دلیل انتخاب

بیش از ۱۰۰۰ کارگذار بیمه‌ی یکی از شرکت‌های بیمه‌گر در ایران می‌باشد.

۷-۵. تعریف سازه‌ها و تعریف عملیاتی متغیرهای پژوهش و نحوه سنجش آنها

تعدادی از متغیرهای مورد نظر به شرح زیر می‌باشند:

- مالکیت: دولتی یا غیر دولتی، اطلاعات نشان می‌دهد شرکتهای دولتی به دلیل در اختیار داشتن امکانات بیشتر و متولی بیمه سطح توقع بیشتری از شرکتهای بیمه دارند و معمولاً سعی می‌کنند از سقف اکثر تعهدات استفاده کنند مانند دندانپزشکی، عینک، فیزیوتراپی
- نوع بیمه‌گذار: بنیادی یا غیر بنیادی از آنجا که سهامدار اصلی این شرکت بنیاد می‌باشد معمولاً شرکتهای بنیادی از تعهدات و پوششهای بیشتری برخوردار هستند و با توجه به داشتن طرح بی نام خسارتهای بیشتری دریافت می‌کنند.

- استان - شهر: ضریب خسارت در شهرهای اصلی با توجه به آگاهی بیمه شده ها و همچنین هزینه های بالای درمانی نسبت به شهرهای کوچک بالاتر است و از طرفی به دلیل عدم فرهنگ بهداشتی در شهرهای کوچک و مهاجرت برای درمان از شهرهای کوچک به شهرهای اصلی بنابراین از این جهت ضریب خسارت بیمه شده هایی که در شهرهای کوچک هستند بالاتر از شهرهای اصلی است، بنابراین این آیتم باید در کنار آیتم های دیگر تحلیل شود.
- تعداد بیمه شده: این آیتم در محاسبه حق بیمه کل موثر است طبق قانون اعداد بزرگ هر چه تعداد بیمه شده ها بیشتر شود ضریب خسارت بالا می رود البته این موضوع به فیلدهای دیگر مانند میانگین سنی و تعداد بیمه شده های زن و مرد نیز وابسته است.

۶-۷. منبع، روش و ابزار گردآوری دادهها (دلیل انتخاب)

منبع دادهها یکی از شرکت های بیمه ای فعال در ایران می باشد.

۷-۷. روش تایید پایایی و روایی ابزار جمع آوری دادهها

جمع آوری داده های پژوهش حاضر با استفاده از پایگاه داده ای موجود شرکت بیمه مورد نظر انجام می شود.

۸-۷. روش تجزیه و تحلیل دادهها (تکنیکها و آزمون های به کار گرفته شده با دلیل انتخاب)

داده ها با استفاده از الگوریتم های داده کاوی دسته بندی می شوند و به منظور مقایسه ی الگوریتم های در نظر گرفته شده از معیارهای دقت، ریکال و پریسیشن استفاده می شود که در ادامه به شرح آن پرداخته خواهد شد:

به طور سنتی معمول ترین معیار مورد استفاده در پیش بینی معیار صحت یا نرخ خطا است (He & Garcia, 2009). یک مدل پایه دسته بندی دو کلاسه را در نظر بگیرید، بطوریکه $\{p, n\}$ برچسب کلاس مثبت درست و منفی درست باشد و $\{Y, N\}$ برچسب کلاس مثبت و منفی پیش بینی شده باشد. بنابراین نمایشی از عملکرد مدل دسته بندی می تواند توسط ماتریس درهم ریختگی^۴ (شکل ۲) نشان داده شود. بر اساس این ماتریس چهار پارامتر زیر قابل تعریف است (Linoff & Berry, 2011).

TP: تعداد عناصر مثبتی که به درستی کلاسه بندی شده اند (مثبت درست).

FP: تعداد عناصر منفی که به نادرستی به عنوان مثبت کلاسه بندی شده اند (مثبت اشتباه).

⁴ Confusion

FN: تعداد عناصر مثبتی که به نادرستی به عنوان منفی کلاسه بندی شده‌اند (منفی اشتباه).

TN: تعداد عناصر منفی که به درستی کلاسه بندی شده‌اند (منفی درست).

		True class	
		p	n
Hypothesis output	Y	TP (True Positives)	FP (False Positives)
	N	FN (False Negatives)	TN (True Negatives)
Column counts:		P_C	N_C

شکل ۲: ماتریس درهم ریختگی

بر مبنای این ماتریس پارامترهای زیادی را می‌توان تعریف نمود که سه شاخص مهم ریکال^۵ و پرسیشن^۶ و صحت^۷ بر اساس این پارامترها تعریف می‌شوند:

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Accuracy} = \frac{TP+TN}{P_C+N_C}$$

⁵ Recall

⁶ Precision

⁷ Accuracy

۸. دستاوردها و نتایج مورد انتظار

دستاورد و نوع استفاده	نام سازمان	ردیف
<p>۱- ایجاد مدلی نو در راستای تحقق اهداف پیشرفت علمی</p> <p>۲- نگارش مقاله‌ای از مدل ایجاد شده و چاپ آن در ژورنالی معتبر</p>	برای دانشکده و دانشگاه	۱
<p>۳- تعیین میزان ریسک بیمه‌گذاران توسط سازمان‌های بیمه‌گر و رفتار متفاوت با بیمه‌گذاران مختلف</p>	برای جامعه و سازمانها	۲
<p>۴- تعیین مشخصه‌ای مهم و تأثیرگذار بر میزان ریسک بیمه‌گذاران</p>	مضامین کاربردی برای مدیران	۳

۹. سهم دانش‌افزایی پژوهش (Contribution and Originality)

تاکنون تحقیقات محدودی به منظور تحلیل رفتار مشتریان در صنعت بیمه در ایران به خصوص بیمه‌نامه درمان با استفاده از روش‌های داده‌کاوی صورت گرفته است، این تحقیق به دنبال پرکردن این شکاف است. به این ترتیب که از طریق داده‌کاوی مشتریان بیمه‌نامه‌های درمان را در یک شرکت بیمه‌ای دسته‌بندی نموده و بهترین مدل را برای دسته‌بندی مشتری و ارائه خدمات مناسب به آن، در این صنعت انتخاب نماید و راهکارهایی را ارائه دهد.

۱۰. زمان‌بندی اجرای پژوهش

نمودار گانت (زمان‌بندی پروژه به ماه)

زمان مورد نیاز												مراحل پژوهش
۱۲	۱۱	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	
												تهیه و تصویب پروپوزال
												بررسی ادبیات پژوهش
												گردآوری داده‌ها
												تحلیل داده‌ها
												نگارش پایان‌نامه

۱۱. تعریف مفهومی واژه‌های کلیدی

- **بیمه:** عملی است که به موجب آن یک طرف (بیمه‌گر) تعهد می‌کند در ازاء دریافت وجه یا جوهی (حق بیمه) از طرف دیگر (بیمه‌گذار) در صورت وقوع یا بروز حادثه برای موردی که بیمه شده، خسارت وارده را جبران نماید و یا وجه معینی را بپردازد (Patrik, 2006).

- **داده کاوی:** عبارت است از فرایند استخراج اطلاعات معتبر، از پیش ناشناخته، قابل فهم و قابل اعتماد از پایگاه داده‌های بزرگ و استفاده از آن در تصمیم‌گیری در فعالیتهای تجاری مهم (Hand, Mannila, & Smyth, 2001).

بیمه درمان: به دلیل بالارفتن هزینه‌های درمان به خصوص در بیماری‌های خاص و به منظور حمایت از افراد جامعه در زمینه جبران هزینه‌های درمانی ناشی از بیماری و حوادث تسهیلات گسترده‌ای توسط بیمه‌گر (شرکت بیمه) برای بیمه‌شدگان به وجود آمده است. مؤسسات، سازمان‌ها و شرکت‌های دولتی و خصوصی که تعداد کارکنان آنها با اعضای خانواده‌هایشان از ۵۰ نفر بیشتر باشد می‌توانند با پرداخت حق بیمه مختصری (به صورت ماهانه، سه ماهه، شش ماهه یا سالانه) تحت پوشش این نوع بیمه قرار گیرند. سقف تعهدات بیمه‌گر در مورد هزینه‌های درمانی با توجه به نیازهای بیمه‌گذار و توافق بین بیمه‌گر و بیمه‌گذار تعیین شده و حق بیمه متناسب با آن محاسبه می‌شود (Y M Chae, 2001).

۱۲. فهرست منابع و مآخذ (بر اساس فرمت APA)

- Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. Cambridge, Mass.: MIT Press.
- He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, 21, 1284.
- Lee, S. J., & Siau, K. (2001). A review of data mining techniques. *Industrial Management & Data Systems*, 101, 41–46.
- Lejeune, M. A. P. M. (2001). Measuring the impact of data mining on churn management. *Internet Research*, 11, 375–387.
- Linoff, G. S., & Berry, M. J. A. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*. Indianapolis, Ind.: Wiley Pub., Inc.
- Patrik, G. (2006). Reinsurance. In *Encyclopedia of Actuarial Science*. John Wiley & Sons, Ltd. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/9780470012505.tar017/abstract>
- Rygielski, C., Wang, J.-C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, 24, 483–502.
- Xie, Y., Li, X., Ngai, E. W. T., & Ying, W. (2009). Customer churn prediction using improved balanced random forests. *Expert Systems with Applications*, 36, 5445–5449.
- Y M Chae, S. H. H. (2001). Data mining approach to policy analysis in a health insurance domain. *International Journal of Medical Informatics*, 62(2-3), 103–111. doi:10.1016/S1386-5056(01)00154-X

۱۲. انطباق موضوع با اولویتهای پژوهشی کشور

۱-۱۳. مرتبط با سند چشم‌انداز (بند و شرح)

۲-۱۳. نقشه جامع علمی کشور (بند و شرح)

۳-۱۳. رسالتهای کلی نظام جمهوری اسلامی ایران / سیاستهای کلی برنامه (بند و شرح)

۴-۱۳. برنامههای توسعه کشور (بند و شرح)

۵-۱۳. سیاستها، برنامهها و دستورالعملهای ابلاغی از وزارت علوم، تحقیقات و فناوری

۶-۱۳. نیازهای دستگاههای اجرایی کشور

۷-۱۳. اولویتهای پژوهشی دانشگاه

امضای دانشجو:

تاریخ: