

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

سمینار

کاربرد داده کاوی در کشف تقلب

نگارنده

چکیده

در سالهای اخیر با افزایش تقلب های مالی در سراسر دنیا و متضرر شدن سازمان ها بر اثر این پیامد، همگان را بر آن داشته است که چگونه می توان راهی برای جلوگیری از آن پیدا کرد. به همین منظور مدیران سازمان ها و بنگاه های مالی به دنبال روشی برای کاهش این ضررها در سازمان های خود می باشند. کشف تقلب روز به روز اهمیت فراوانی پیدا کرده و به یک مسئله پیچیده و سخت در محیط های صنعتی امروز تبدیل شده است. تکنیک های داده کاوی می توانند روش های مؤثری برای حل این مشکل عمده باشند. در این سمینار، با معرفی داده کاوی، به اهمیت آن در زمینه کشف تقلب های مالی پرداخته شده است.

واژه های کلیدی: داده کاوی، تقلب، حملات، مالی.

فهرست مطالب

۸	فصل ۱: مقدمه
۹	۱-۱- مقدمه
Error! Bookmark not defined.	فصل ۲: تقلب و بررسی آن
۵	۱-۲- جایگاه و ضرورت تشخیص تقلب در بیمه
۵	۲-۲- مفهوم تقلب بیمه ای
۶	۳-۲- تقلب و متقلبان
۸	۴-۲- بررسی تقلب در صنعت بیمه
۹	۵-۲- مشخصات انواع کلاهبرداری و تقلب در صنعت بیمه
Error! Bookmark not defined.	فصل ۳: داده کاوی
۱۱	۱-۳- مقدمه
۱۱	۲-۳- تعاریف داده کاوی
۱۴	۳-۳- معرفی داده کاوی
۱۴	۴-۳- مفهوم داده کاوی
۱۵	۵-۳- هدف داده کاوی
۱۵	۶-۳- داده کاوی و سایر رشته ها
۱۶	۷-۳- آمار و داده کاوی
۱۹	۸-۳- ضرورت استفاده از داده کاوی
۲۰	۹-۳- وظایف داده کاوی
۲۳	۱۰-۳- ساختار یک سیستم داده کاوی
۲۶	۱۱-۳- فرایند داده کاوی
۲۷	۱۲-۳- فعالیتهای مرتبط با پژوهش
۳۱	۱۳-۳- فنون داده کاوی
۳۲	۱۴-۳- روشهای داده کاوی مورد استفاده در کشف تقلبهای مالی
۳۴	۱۵-۳- مدل ها و الگوریتم های داده کاوی
۳۹	۱۶-۳- مشکلات داده کاوی
۴۰	۱۷-۳- محدودیت ها
۴۰	۱۵-۳- کاربرد داده کاوی

فصل ۴: جمع‌بندی و پیشنهادها

- ۴-۱- نمونه‌های اجرا شده داده کاوی ۴۵
- ۴-۲- داده کاوی در ایران ۴۵
- ۴-۳- ضرورت و جایگاه داده کاوی در تامین اجتماعی ۴۶
- ۴-۴- ملزومات اجرای داده کاوی ۴۸
- پیشنهادات ۴۹

Error! Bookmark not defined.

مراجع

فهرست اشکال

- شکل ۳-۱- معماری سیستم داده کاوی..... ۲۴
- شکل ۳-۲- دسته بندی تحقیقات در زمینه داده کاوی..... ۲۸
- شکل ۳-۳- روشهای داده کاوی استفاده شده در کشف انواع تقلب..... ۳۳

فهرست جداول

- جدول ۱-۲- طبقه بندی تقلبهای مالی ۶
- جدول ۱-۳- مقایسه آمار و داده کاوی ۱۷

فصل اول

مقدمه

دنیای مدرن در واقع دنیایی داده‌گرا است. ما با داده‌های عددی و غیر عددی در بخش های صنعتی، بازرگانی، مدیریتی، مالی و علمی احاطه شده‌ایم. این داده‌ها باید تحلیل و پردازش شوند تا تبدیل به اطلاعاتی شوند که آگاهی بخش، آموزش دهنده و پاسخگو باشند یا به شکل دیگری به درک و تصمیم‌گیری ما کمک کنند. در این راستا تقلب و سوء استفاده های مالی نیز به موازات رشد تکنولوژی و داده‌های آماری گسترش پیدا کرده است. تقلب یک تجارت میلیون دلاری هست که در سالهای اخیر روند صعودی را داشته است. در سال ۲۰۰۴ و ۲۰۰۵ حدود ۴۵٪ از شرکتهای و کمپانیهای سراسر دنیا به جهت تقلب و سوء استفاده های مالی متضرر شده اند بطوریکه بر اساس گزارش سازمان جهانی جرایم اقتصادی و به استناد به فاکتورهای محدودی از جمله (اظهار غیر واقعی و جعل) در سال ۲۰۰۵ این خسارت بالغ بر ۱,۷ میلیون دلار آمریکا اعلام شده است. نکته قابل تامل اینست که ۳۴٪ این تقلبات بصورت تصادفی کشف شده است و این خود اهمیت و ضرورت وجود یک سیستم منسجم و کارآمد جهت پیشگیری و کشف تقلب را یادآوری می نماید. کشف تقلب راههای مختلفی دارد و تمیزی داخلی یکی از روشهای سنتی میباشد که در این روش ممیز و کارشناس داخلی اقدام به پیشگیری و کشف تقلب می کند ولی بدلیل نا کارآمد بودن این روش سیستم های پیشرفته ای از قبیل ممیزی خارجی، سیستم مدیریت خطر، سیستم ایمنی اطلاعات (ایزو ۲۷۰۰۰) و داده کاوی^۱ جهت پیشگیری و کشف تقلب معرفی گردیده اند. در سالهای اخیر رشد انفجار آمیزی از روش ها برای اکتشاف دانش جدید از داده‌های خام به وقوع پیوسته است. در پاسخ به این امر، رشته‌ی جدیدی از کاوش داده‌ها موسوم به داده‌کاوی به طور ویژه گسترش یافته است تا اطلاعات با ارزشی از مجموعه داده‌های عظیم استخراج نمایند.

امروزه فناوری داده‌کاوی به صورت موضوعی داغ برای تصمیم‌گیران در آمده است، زیرا این فن، اطلاعات نهفته ی با ارزش تجاری و علمی را از داده‌های حجیم ذخیره شده استخراج می‌کند. با این وجود، ماهیت داده‌کاوی فناوری جدیدی محسوب نمی‌شود. استخراج اطلاعات و دانش از داده‌های ذخیره شده یک مفهوم کاملاً دیرینه در مطالعات علمی و پزشکی می‌باشد. آنچه که جدید است همگرایی و اشتراک چندین رشته و فناوری‌های متناظر آنها است که فرصت منحصر به فردی برای داده‌کاوی به دنیای علم، تجارت و اقتصاد معرفی کرده است. اما داده‌کاوی یا کشف دانش و معرفت از پایگاه داده‌ها با این تعریف یک شاخه‌ی نسبتاً جدید علمی است که از انجام

تحقیقات در رشته های آمار ، یادگیری ماشین^۱ ، علوم رایانه (به ویژه پایگاه داده ها) شکل گرفته است. در حالی که مرزهای این رشته ها در داده کاوی مبهم است، ولی می توان گفت که مهمترین این رشته ها آمار می باشد به طوری که بدون آمار داده کاوی مفهومی نخواهد داشت. با وجود اینکه داده کاوی یک رشته ی نسبتاً جدید علمی می باشد و کاربردهای متنوع و گسترده ای در رشته هایی مانند بازرگانی، پزشکی، مهندسی، علوم رایانه، صنعت، کنترل کیفیت، ارتباطات و کشاورزی پیدا کرده است. امروزه مرز و محدودیتی برای دانش داده کاوی متصور نیست و مرز آن را از اعماق اقیانوس ها تا بیکران فضا می دانند. به عبارت دیگر کاربرد دانش داده کاوی در تمامی حوزه های برخوردار از داده بوده و تنها محدودیت دانش داده کاوی نبود داده است.