



کنگره بین المللی علوم و مهندسی

آلمان - هامبورگ

اسفند ماه ۱۳۹۶

تشخیص تقلب‌های آنلاین بانکی در داده‌های بسیار نامتوازن با

استفاده از داده‌کاوی (مطالعه موردی)

مریم صدقی مرادی^۱، حسن نادری^۲

۱- دانشجوی کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی قزوین،

sedghi.moradi@outlook.com

۲- استادیار، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی قزوین، naderi@iust.ac.ir

چکیده

در طول دهه‌های اخیر حجم زیادی از داده‌ها در پایگاه داده‌ها ذخیره شده‌اند. در واقع سازمان‌ها علیرغم نیاز به اطلاعات نهفته در داده‌ها، قادر به استفاده از آنها نیستند. از طرف دیگر، با توجه به تنوع زیاد مشتریان، بازارها و تنوع و پیچیدگی محیط‌های کسب و کار، لزوم استفاده از راهکارهای مناسب برای طبقه‌بندی و یافتن اطلاعات کاربردی جهت تصمیم‌گیری صحیح و به موقع در میان انبوهی از داده‌ها برای سازمان‌ها امری ضروری و حیاتی می‌باشد. در این میان، روند رو به افزایش تقلب‌های بانکی، منجر به کاهش اعتبار این موسسات در برآوردن انتظارات مشتریان خود می‌شود. معذک، به نظر می‌رسد تحلیل و آنالیز اطلاعات حاصل از تراکنش‌های موسسات مالی و به تبع آن تدوین سیاست‌ها و پروتکل‌های امنیتی و احراز هویت امری غیر قابل اجتناب باشد.

در این تحقیق پایه اصلی کار براساس متدولوژی استاندارد داده‌کاوی ^۱CRISP-DM بنا شده‌است. براساس متدولوژی استاندارد CRISP-DM، ابتدا مسائل اصلی تحقیق و فضای کلی کار را بررسی نموده و اقدام به تهیه مجموعه داده مرتبط با مسئله نموده ایم. سپس، با استفاده از روش‌های آماری و جداول و نمودارها تحلیل‌های اولیه را بر روی داده‌ها انجام داده‌ایم. در ادامه به آماده‌سازی و پیش‌پردازش داده‌ها پرداخته شده‌است. در طول اجرای مرحله آماده‌سازی، تعدادی از مشخصه‌ها ایجاد، حذف یا تغییر می‌یابند. در مرحله چهارم، مدل‌سازی را با استفاده از تکنیک‌های متفاوت رده‌بندی انجام می‌شود. ارزیابی مدل‌های ایجاد شده در آخرین مرحله از تحقیق انجام شده‌است که به دنبال آن مدل‌های بهینه‌سازی دقیق‌تر را به عنوان مدل نهایی مسئله انتخاب شده‌است. در مرحله ارزیابی، نتایج به دست آمده را با مسئله مورد بحث، که در قالب سوالات اصلی تحقیق معرفی می‌شود، مقایسه نموده و میزان نیل به اهداف تحقیق را مورد سنجش قرار داده می‌شود.

واژه‌های کلیدی: کشف تقلب، داده‌کاوی، داده‌های حجیم، متدولوژی استاندارد CRISP-DM

¹Cross-industry Standard Process for Data Mining