

استفاده از روش آنسامبل ترکیبی برای تحلیل داده‌های مشتریان بیمه بدنه بیمه ملت

هادی جعفری^۱، یاسر صمیمی^۲

^۱کارشناس ارشد، دانشگاه خواجه نصیرالدین طوسی؛ hadi.jafari.dm@gmail.com

^۲استادیار، دانشگاه خواجه نصیرالدین طوسی؛ y_samimi@kntu.ac.ir

چکیده

یکی از تصمیمات مهم در موسسات مالی ارزیابی اعطای تسهیلات و کاهش ریسک اعتباری است. حال با توجه به حجم عظیم داده‌های اعتباری مشتریان، داده‌کاوی می‌تواند جهت دسته‌بندی و شناسایی مشتریان خوش حساب و بدحساب مورد استفاده قرار گیرد. اما در بیشتر موارد، داده‌های اعتباری از مشکل عدم توازن رنج می‌برند که موجب شده است اکثر الگوریتم‌های داده‌کاوی رکوردهای دسته اقلیت را نادیده بگیرند. برای حل این مشکل، یک مدل آنسامبل ترکیبی استفاده و بر روی مجموعه داده بیمه ملت پیاده‌سازی شد. عدم قابل فهم بودن مدل بدست آمده مشکلی در روش آنسامبل است که در این تحقیق از یک روش استخراج قانون برای قابل فهم کردن نتایج تولید شده بهره گرفته شد. نتایج بیانگر این است که مشتری جدیدی که شامل تخفیف بالای ۱۷.۸٪ می‌شود با احتمال زیاد در طول یک سال خسارتی را متوجه شرکت بیمه نخواهد کرد.

کلمات کلیدی

امتیازدهی اعتباری، داده‌کاوی، دسته‌بندی، عدم توازن، شرکت بیمه ملت، روش آنسامبل، استخراج قانون

Using a hybrid ensemble method for analyzing customers data of Mellat insurance company

Author's Name, Author's Name

Author's affiliations

ABSTRACT

One of the most important factors in decision making about providing financial services in financial institutions is customer's level of credit risk. Nowadays, with the accessibility of the huge amount of data of customers' creditworthiness, data mining methods could be used for the purpose of risk-based classification of customers. Nevertheless, in most of the cases, credit data suffer from imbalance problem and as a result, data mining algorithms ignore records from minority class. To address this problem, a hybrid ensemble method was used and implemented on customers data of Mellat insurance company. An ensemble model is not understandable. Due to that, a rule extraction method was used. Consequences show that a new customer who is qualified for a discount higher than 17.8% most likely will not claim for any money.

KEYWORDS

credit scoring, data mining, classification, imbalance, Mellat insurance company, ensemble method, rule extraction

^۱ هادی جعفری؛ میدان ونک، خیابان ملاصدرا، خیابان پردیس، دانشکده مهندسی صنایع دانشگاه خواجه نصیرالدین طوسی؛ ۰۹۱۲۷۹۶۷۱۱۲

این مقاله از پایان‌نامه‌ای استخراج شده که با همکاری و حمایت پژوهشکده بیمه انجام پذیرفته است