A Study on Validity Detection for Shipping Decision in the Mail-order Industry

Masakazu Takahashi\textsuperscript{a*}, Hiroaki Azuma\textsuperscript{b}, and Kazuhiko Tsuda\textsuperscript{c}

\textsuperscript{a}Graduate School of Innovation and Technology Management, Yamaguchi University, 2-16-1 Tokiwa-dai, Ube, Yamaguchi 755-8611, Japan
\textsuperscript{b}HAZS Corporation, 6-2-6, Kojinmachi, Chiyoda, Tokyo 102-0083, Japan
\textsuperscript{c}Graduate School of Business Sciences, University of Tsukuba, 3-29-1, Otsuka, Bunkyo, Tokyo 112-0012, Japan

Abstract

This paper presents investigating fraud transaction detection in the mail order industry. These kinds of detection have done intensively, but the outcome of the research has not shared among the mail-order industry. As the B2C market such as the Amazon type business expands their market volume exponentially, the fraud transactions increase in number. As a matter of course, this phenomenon is not only continuing but clever. One of the conclusive factor for this phenomenon is the payment method. That is, the deferred payment method. The conventional primary indicator for the fraud detection is the ordered time based information. They are shipping address, recipient name, and the payment method. This kind of information makes use of the prediction in common. Conventional detecting method for the fraud depends on the human working experiences so far. From such kind of information, the mail-order company predicts the potential fraud customer with their working experience parameters. As the number of order transaction becomes large, fraud detection becomes difficult. The mail order industry needs something clever detection method. From these backgrounds, we observe the transaction data with the customer attribute information gathered from a mail order company in Japan and characterized the customer with a machine learning method. From the results of the intensive research, potential fraudulent transactions are identified. Intensive research revealed that the classification of the deliberate customer and the careless customer with machine learning.

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of KES International

Keywords: Risk Management; Mail Order; Customer Analyses; Fraud Detection; Machine Learning; Weak Learner; Random Forest; Service Science and Management Engineering

\textsuperscript{*} Corresponding author. Tel.: +81-836-85-9066; fax: +81-836-85-9877.
E-mail address: masakazu@yamaguchi-u.ac.jp
1. Introduction

Mail-order industry is a typical non-face-to-face sales system and to provide such goods or services to the customer’s hand directly via logistics company. A post-paid system method is popular among the domestic mail-order company and this payment method provides the effect of reducing customer’s anxiety. However, some customers with malicious who do not pay the price while receiving the products are existed along with the mail-order industry market size expands year by year.

A survey from the Japan Direct Marketing Association [1] says the percentage of credit losses of a mail order company is estimated about 0.5% of net sales. Therefore, it is important for the mail order company to predict risk exposures in customers’ credit control. It is because too large risk exposure leads to high default risk and too small risk exposure misses their business opportunities. Therefore, for the mail order industry, the credit scoring strategy is a serious problem for the further market expansion.

The difference between the mail-order and the onsite is the distance and time gap among the customers and store. Mail-order company cannot see the customer, on the other hand, the customer also cannot see the mail-order company through. Among the mail-order industry, it is in fact that the malicious company exists among the industry. Therefore, there are many customers who do not promise the mail-order company, as a result.

In such a conflicting issue of the credibility, some effective method is required among the mail-order industry. Recently, a big data analysis is in fashion. This is one of the prediction method making use of huge amount of data. This method is adopted widely to many fields. But, the result of the Google Flu Trends shows, an accurate prediction needs the heuristics knowledge. For these backgrounds, we observe the transaction data with the customer attribute information gathered from a mail order company in Japan and characterized the customer with machine learning method.

In this paper, we carry out data analysis for the customer behavior through the bad debt list in a mail order company. The rest of the paper is organized as follows: Section 2 discusses the backgrounds of the research and related works; Section 3 briefly summarize the gathered data on the target mail order company; Section 4 describes analytics of the data and presents analytical results; and Section 5 gives some concluding remarks and future works.

2. Background and Related work

First, we describe the research objectives through the current condition for the mail-order industry. As the market size of the mail order industry grow, the more fraud transaction increased. This is because the mail-order company must process the received order year by year. Furthermore, a post-pay system, one of the characteristics of the domestic mail-order company is the primary reason for the fraud transaction. Then, about 20% of the delivery packets fails into the re-scheduled delivery [2]. Since not only the mail-order company but the logistics company is fed up the daily transaction, drastic measures against fraud have not been taken, so far. Therefore, we intend to derive the heuristics knowledge which is the basis of our proposed method from data automatically collected. Next, we describe the potential of Big data.

Decision-making using data becomes a primary business activity since the keywords of big data is in common, recently. Another key word such as “IoT” stands Internet of Things, also contribute the deep understanding of the data oriented decision making. As a result, the service from the data analysis result is provided in various phase on our life.

In this section, we describe the research backgrounds, especially focusing on the merits and demerits of big data, then describe the research trends including the surroundings of the proposed method. At first, describing the side effect of big data analysis.

In 2009, Google announced the GFT (Google Flu Trends) from the analytical result of their search engine [3]. Schnberger described the algorithm of the GFT [4]. Google assembled the algorithm with the five years’ web logs including the hundred billion of search results and then proposed the prediction influenza indicators with the 45 search words. They explained the model more effective than the government alert from the statistical data analysis that delays announce. However, the GFT was not able to predict the outbreak of the pig flu in 2009. The GFT predicted the epidemic of influenza excessively about 1.5 times occurred in the end of 2012 [5]. Moreover, the
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات