A Data Mining Model for Risk Assessment and Customer Segmentation in the Insurance Industry

Payam Hanafizadeh, Department of Management and Accounting, Allemeh Tabataba'i University of Tehran, Iran

Neda Rastkhiz Paydar, Department of Management and Accounting, Allemeh Tabataba'i University of Tehran, Iran

ABSTRACT

Customer segmentation on the basis of predictable risks can help insurance firms maximize their earnings and minimize their losses. Car insurance is one of the most lucrative and profitable branches in the insurance industry. Utilizing the concept of self-organizing map, the authors propose a two-phase model called 'Auto Insurance Customers Segmentation Intelligent Tool' to segment customers in insurance companies on basis of risk. In the first phase, the authors extract 18 risk factors in four categories consisting of demographic specifications, auto specifications, policy specifications, and the driver's record extracted from the literature review. In the second phase, they finalize the selection process by drawing on expert opinion polls. The authors utilize self-organizing maps since they are able to display the output in the form of illustrative and comprehensible graphical maps capable of representing linear and non-linear relationships among variables, insensitive to the learning input, and slightly sensitive to the noise in the learning input. Finally, K-means are employed to compare the results with those obtained through self-organizing maps.

Keywords: Comprehensive Auto Insurance, Customer Segmentation, K-means, Self-organizing Map

1. INTRODUCTION

Currently, one of the most essential economic sectors in every country is the insurance industry that fulfills vital duties and functions. The growth of insurance industry in every country is an indicator of the development and the rise of financial savings. This crucial industry, which offers appropriate services, promotes people's welfare in a community (Jaafary, Samimi, & Morady, 2008). Since the customer is considered as the fundamental factor in producing income in businesses and raising profitability, how to interact with the customer determines the amount of profit. In addition, segmentation has been identified as one of the most significant concepts in marketing. Skillful segmentation helps firms to pick out profitable customers, un-

DOI: 10.4018/jsds.2013010104

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

derstand customers' desires, allocate resources, and support them against competitors (Dickson, 1982). Customers vary in many essential ways. Generally, customers can be distinguished from each other from two different views, that is, from their value for the organization, and the difference in their needs (Hosseinzadeh, 2008). Cycle of a customer relationship management (CRM) consists of four dimensions. CRM begins with customer identification. This phase involves targeting the population that is most likely to become customers or most profitable to the company. Elements for customer identification include target customer analysis and customer segmentation. Target customer analysis involves seeking the profitable segments of customers through analysis of customers' underlying characteristics, whereas customer segmentation involves the subdivision of an entire customer base into smaller customer groups or segments, consisting of customers who are relatively similar within each specific segment (Ngai et al., 2009). Risk segmentation is, in fact, the segmentation of customers with similar risk characteristics, probably causing similar damage. Risk segmentation of policyholders on the basis of observable features can help insurance companies to reduce loss, raise the rate of insurance coverage, and prevent them from making an inappropriate choice in the insurance market. The majority of developed and developing countries take advantage of the risk segmentation in determining the insurance premium in their insurance industry (Majed, 2008). The rate of the insurance premium in many insurance companies is computed with regard to different demographic variables, car specifications, and the record of damage caused by the car owner. Mike Kreidler, a member of the board of directors of one of the insurance companies in Washington State, pointed out that the rate of insurance premium relies on factors such as the policyholder's age, gender, and marital status, vehicle type, the location of the car owner's residence, the claim history, and the driving pattern (Kreidler, 2008). While in Iran the rate of comprehensive auto insurance premium is set by Central Insurance of Iran. In practice, low-risk customers pay for the damage and loss caused by high-risk customers, so there is no difference between these two groups of customers. In addition to the inefficiency of the contracts of insurance or policies, lack of such measures in determining the risk in automobile insurance leads to computing unfair rates because in these cases instead of the person, the car is insured. That's why most insurance companies experience great loss as far as automobile insurance is concerned while most developed countries have attempted to increase the productivity and profitability of their insurance industry using the risk segmentation system (Hosseinzadeh, 2008).

Self-organizing map (SOM) is a type of artificial neural network that is trained using unsupervised learning to produce a representation of the output in the form of visual graphical maps which are comprehensible to managers of organizations. The insensitivity of the self-organizing map to the learning input, and its slight sensitivity to the noise in the learning input, its capability to display linear and nonlinear relationships among variables, and its great power to segment data are among the advantages of these maps. In this research, this map in data mining is used to segment customers of automobile comprehensive insurance using the recognized factors in risk. Since this map is able to illustrate the output in the form of graphical maps, it is easier for managers and experts to understand and interpret the results.

In the present study, the important and effective factors in the risk of policyholders will be identified and picked out in two separate phases. In the first phase, the study and review of the articles published in the prestigious journals including 'Science direct', 'IEEE', 'Emerald', 'ProQuest' and scientific reports within a span of 9 years from 2000-2009 was carried out to select risk factors. The extracted factors were generally grouped into four main categories, that is, demographic specifications, automobile specifications, insurance policy terms, and the driver's record. In the second phase, a questionnaire concerning the identified factors was designed and filled out by experts from among those working in financial compensation department and automobile comprehensive

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/data-mining-model-risk-assessment/77336

Related Content

Systems Thinking and Cognitive Process in Marketing
(2012). *Systems Thinking and Process Dynamics for Marketing Systems: Technologies and Applications for Decision Management (pp. 170-197).*
www.irma-international.org/chapter/systems-thinking-cognitive-process-marketing/65306

Putting Context in the Network Access of Mobile Applications Using Fuzzy Analytic Hierarchy Process
Yaser A. Mowafi, Tareq Alaqarbehand Rami Alazrai (2019). *International Journal of Decision Support System Technology (pp. 13-26).*
www.irma-international.org/article/putting-context-in-the-network-access-of-mobile-applications-using-fuzzy-analytic-hierarchy-process/223425

Comparison of Normalization Techniques on Data Sets With Outliers
Nazanin Vafaei, Rita A. Ribeiroand Luis M. Camarinha-Matos (2022). *International Journal of Decision Support System Technology (pp. 1-17).*
www.irma-international.org/article/comparison-of-normalization-techniques-on-data-sets-with-outliers/286184

Metamodel of the Artifact-Centric Approach to Event Log Extraction from ERP Systems
Ana Pajiand Dragana Beejski-Vujaklija (2016). *International Journal of Decision Support System Technology (pp. 18-28).*
www.irma-international.org/article/metamodel-of-the-artifact-centric-approach-to-event-log-extraction-from-erp-systems/157363

Development and Application of a Spreadsheet-Based Spatial Decision Support System (SDSS)
Dossay Oryspayev, Ramanathan Sugumaranand John DeGroote (2011). *International Journal of Decision Support System Technology (pp. 1-14).*
www.irma-international.org/article/development-application-spreadsheet-based-spatial/62563