Chapter 77
Data Mining and Explorative Multivariate Data Analysis for Customer Satisfaction Study

Rosaria Lombardo
Second University of Naples, Italy

ABSTRACT

By the early 1990s, the term "data mining" had come to mean the process of finding information in large data sets. In the framework of the Total Quality Management, earlier studies have suggested that enterprises could harness the predictive power of Learning Management System (LMS) data to develop reporting tools that identify at-risk customers/consumers and allow for more timely interventions (Macfadyen & Dawson, 2009). The Learning Management System data and the subsequent Customer Interaction System data can help to provide "early warning system data" for risk detection in enterprises. This chapter confirms and extends this proposition by providing data from an international research project investigating on customer satisfaction in services to persons of public utility, like education, training services and health care services, by means of explorative multivariate data analysis tools as Ordered Multiple Correspondence Analysis, Boosting regression, Partial Least Squares regression and its generalizations.

INTRODUCTION

The necessity of systems of evaluation and assessment in many socio-economic fields together with the Learning Management System data (LMSD) have sparked a need in building early warning system (EWS) which produces signal for possible risks. Accordingly various EWSs have been established (Kim *et al.*, 2004): for detecting fraud, for credit-risk evaluation in the domain of financial analysis (Phua, *et al.*, 2009), to detection of risks potentially existing in medical organizations (risk aversion of nurse incidents, infection control and hospital management), to support decision making in customer-centric planning tasks (Lessman & Vob, 2009).

DOI: 10.4018/978-1-4666-2455-9.ch077

Enterprises are implementing systems of evaluation and assessment to demonstrate their commitment to efficiency, productivity, effectiveness and accountability (Volkwein, 1999). Managers are therefore under pressure to implement and demonstrate effective best practice. At the same time, their own performance as managers is under increasing government, public, and customers/ consumers scrutiny.

In this chapter we focus on EWS of LMSD for customer-centric planning tasks, to develop exploratory tools that identify at-risk customers and allow for more timely interventions.

Several data mining classifiers/predictors will be probed as training tools for detecting risks coming from Customer Interaction System data. The "Customer Interaction System" collects different information on customers/consumers by database and data-warehouse dealing with customers handled by the Consumer Affairs and Customer Relations contact centers within a company.

The maximization of lifetime values of the customer base in the context of a company's strategy is a key objective of customer relationship management (CRM). The role of CRM is in supporting customer-related strategic measures. Customer understanding is the core of CRM. It is the basis for maximizing customer lifetime value, which in turn encompasses customer segmentation and actions to maximize customer conversion, retention, loyalty and profitability. Hence, emphasis should be put on correct customer understanding and concerted actions derived from it, in order to reduce the risk to lose customers by loyalty increase.

The main aim of this project is to focus on customer satisfaction in services to persons of public utility, like training services and health care services.

The training services and the health care ones, as services of utility towards the public, are of relevant social interest. The main result of the supply process is not an output external to people, but it identifies itself with the effects of the service on consumers. For this reason the "product" is not simply an *output* (as result of a short period), but an *outcome* (as result of a long period) which, in health care services, is given by the psychophysical state or condition of health of consumers/customers who perceived the service, while in training services, it can be identified with the success to find appropriate works (Goldstein & Spiegelhalter, 1996; Gori & Vittadini, 1999). The quality of health care services (Drummond *et al*., 1997; Donabedian, 1985), supplied by a plurality of agents to consumer people, should be measured not only by quantitative-physic indicators (number of beds, number of operating rooms, number of high technologic instruments, number of training courses, number of teachers, etc.) and quantitative-monetary ones (input/output ratios, productivity indexes, etc.), but also by qualitative indicators which allow to measure the customer satisfactions, based essentially on the subjective evaluation of the quality service.

The availability of *customer satisfaction* measures is becoming as important as the corresponding objective measures, so it represents an early warning system data set, one of the essential component of the "optimal business management" following the approach of the Total Quality Management. The more reliable models for quality of services developed within a scientific framework agree on the necessity of a precise recognition of the fundamental perceptive dimensions on which the customer bases his or her judgment of the quality of particular public services. This evaluation of quality is typically express using ordinal categorical variables, which often require suitable transformation functions. Throughout this chapter with respect to different statistical methods, we consider different transformation functions of the ordinal categorical variables.

For the study of customer satisfaction, the theoretical models propose a system of five aspects of quality, these are: *tangibility, reliability, capacity of response, capacity of assurance and empathy*. These various aspects of the quality of