Expert Systems with Applications 39 (2012) 12772–12783

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

On the differential benchmarking of promotional efficiency with machine learning modeling (I): Principles and statistical comparison

Cristina Soguero-Ruiz [a], Francisco-Javier Gimeno-Blanes [b], Inmaculada Mora-Jiménez [a], María Pilar Martínez-Ruiz [c], José-Luis Rojo-Álvarez [a,*]

[a] Signal Theory and Communications Department, University Rey Juan Carlos, Camino del Molino s/n, Fuenlabrada, 28943 Madrid, Spain
[b] Signal Theory and Communications Department, University Miguel Hernández, Av. Universidad s/n, Elche, 03202 Alicante, Spain
[c] Commercialization and Market Research Department, University Castilla La Mancha, Av. de los Alfares, 44, 16071 Cuenca, Spain

ARTICLE INFO

Keywords:
Promotion
Machine learning
Bootstrap
Price indices
Marketing
Signal processing

ABSTRACT

Sales promotions have become in recent years a paramount issue in the marketing strategies of many companies, and they have even more relevance in the present economic situation. Currently, the empirical models, aimed at assessing consumers behavior in response to certain sales promotions activities such as temporary price reductions, are receiving growing attention in this relevant research field, due to two reasons mainly: (1) the complexity of the interactions among the different elements incorporated inside promotions campaigns attracts growing attention; and (2) the increased availability of electronic records on sales history. Hence, it will become important that the performance description and comparison among all available machine learning promotion models, as well as their design parameters selection, will be performed using a robust and statistically rigorous procedure, while keeping functionality and usefulness. In this paper, we first propose a simple nonparametric statistical tool, based on the paired bootstrap resampling, to allow an operative result comparison among different learning-from-samples promotional models. Secondly, we use the bootstrap statistical description to evaluate the models in terms of average and scatter measurements, for a more complete efficiency characterization of the promotional sales models. These statistical characterizations allow us to readily work with the distribution of the actual risk, in order to avoid overoptimistic performance evaluation in the machine learning based models. We also present the analysis performed to determinate whether the figure of merit has a significant impact on final result, together with an in depth design parameter selection to optimize final results during the promotion evaluation using statistical learning techniques. No significant difference was obtained in terms of figure of merit choice, and Mean Absolute Error was selected for performance measurement. As a summary, the applied technique allows clarifying the design of the promotional sales models for a real database (milk category), according to the influence of the figure of merit used for design parameters selection, showing the robustness of the machine learning techniques in this setting. Results obtained in this paper will be subsequently applied, and presented in the companion paper, devoted to a more detailed quality analysis, to evaluate four well-known machine learning algorithms in real databases for two categories with different promotional behavior.

1. Introduction

The current economic landscape, characterized by financial instability and the consequent changes in consumer behavior, is driving a transformation in food retailer decision, bringing to a new and more aggressive promotional perspective (Quelch, 2008). As an example of this situation, the dramatic sales reduction of food products in Spain, which has led retailers in the industry to implement new approaches, such as the intense use of private label products, can be mentioned. In addition, it has been also searched to increase consumer's frequent purchases through promotional activities, such as promotional discounts, feature advertising, and promotional packs (e.g., "buy 3 get 1 free") (Quelch, 2008). Therefore, there is no denying that sales promotion has become in recent years a key tool for marketing strategies in retail food markets, and for this reason, investment in this area has strongly increased, reaching values over 50% of marketing budgets in relation to other communications tools (Villalba & Iñaki, 2002).

The present economic situation, along with food retailer's strategies and increasing investment in promotional activities, has motivated an important number of research efforts to

* Corresponding author.
E-mail addresses: cristina.soguero@urjc.es (C. Soguero-Ruiz), javier.gimeno@umh.es (F.-J. Gimeno-Blanes), inmaculada.mora@urjc.es (I. Mora-Jiménez), mariapilar.martinez@uclm.es (M.P. Martínez-Ruiz), joseluis.rojo@urjc.es (J.-L. Rojo-Álvarez).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.04.017

characterize sales promotion and to measure promotional efficiency. Existing models for analysing sales promotions effects can be classified into two separate groups. In the first group, namely *theoretical models*, consumer behavior is basically evaluated considering a sociological and psychological perspective, whilst in the second group of *empirical models*, promotional structures based on empirical information extracted from historical databases are usually built.

Within that last group, the efforts have been focused during the last decades on the understanding of sales promotion dynamics based on classical statistical analysis methods, and more recent works are concentrated towards the machine learning algorithmic and data mining techniques, as powerful tools to extract information from existing recorded data (Mitchell, 1997; Van Heerde, Leeflang, & Wittink, 2000). Machine learning techniques aim to find recurring patterns, trends, or rules, which can explain the data behavior in a given context, and then allows to extract new knowledge on the consumer behavior, to improve the performance of marketing operations, and to estimate the commonly called Deal Effect Curve (DEC). In particular, a vast amount of knowledge has been extracted from machine learning techniques, although not all promotional behaviors have been studied and there is still room for in depth further studies (Bell, Chiang, & Padamanabhan, 1999; Blatterg, Briescha, & Fox, 1995; Leeflang & Wittingk, 2000). More specifically, operational problems arise in machine learning promotional modeling, when based on nonlinear estimation techniques, for evaluating and demonstrating working hypothesis (Liu, Kong, & Yang, 2004; Martínez-Ruiz, Mollá-Descals, & Rojo-Álvarez, 2005, 2006b, 2006a; Van Heerde, Leeflang, & Wittink, 2001; Wang, Li, & Zhao, 2008). In this paper we present, prior to an in detail analysis of machine-learning performance, a pre-evaluation of different figures of merit, to asses their impact on the final result, together with an in depth analysis for design parameters selection. Results obtained here allowed establishing and validating figures of merit and the design parameters selection procedure to be applied for pricing promotion study. Specifically, we applied these results in the companion paper (Soguero-Ruiz, Gimeno-Blanes, Mora-Jiménez, Martınez-Ruiz, & Rojo-Álvarez, 2012), to evaluate four well-known machine learning algorithms in two real databases for two categories presenting dramatically different promotional behavior.

The draw of the paper is as follows. Section 2 includes a review of basic concepts of sales promotion in retailer environments, as well as a summary of previous work on machine learning in the context of marketing. Section 3 gives a short description of non-parametric inference paired hypothesis tests, based on bootstrap resampling, and actual risk evaluation of a set of adequate figures of merit is introduced. The figure of merit benchmark represents a relevant contribution of this paper, as it ends up becoming an operative tool for decision support in promotional modeling with machine learning techniques. Section 4 summarizes experiments and results based on real data. Finally in Section 5, conclusion and remarks are presented.

2. Background

Though many definitions have been published for the term *sales promotion* (Blattberg & Neslin, 1990; Kotler & Keller, 2005; Yeshin, 2006), none of them are generally accepted, but general consensus suggests that sales promotions consist basically of short-time sales incentives (Blattberg & Neslin, 1990; Kotler & Keller, 2005). For instance, the American Marketing Association (AMA) defines *sales promotion* as a *media and non media marketing pressure applied for a predetermined, limited period of time in order to stimulate trial, increase consumer demand, or improve product availability* (Goodstein, 2000). Some researches (Blattberg & Neslin, 1990) consider sales promotion not just a marketing element, but instead included within the strategic activity undertaken by the company.

The sales promotion strategy adopted by the grocery retailer must be consistent with the general pricing policy. In fact, some strategic aspects of the retailer's pricing policy cover certain considerations related to the appropriateness of the use of promotions and discounts. For this reason, when under certain circumstances the use of deals and discounts are considered adequate, the specific discount rate must be determined attending to timing, frequency, and magnitude of the promotional discounts, (Krishna, 1994; Kumar & Pereira, 1997; Shankar & Bolton, 2004). There are three main elements which need to be set over a promotion by the retailer:

- First, the strategy adopted by the retailer in relation to price changes. This strategy is essential as it contributes to determine the retailer's strategic positioning in the long term. It can range from maintaining a stable pricing level, with little changes, (i.e. Everyday Low Pricing strategy from Wal-Mart and Mercadona), to intensive price changes and promotions (i.e. High–Low pricing strategies) (Hoch, Dreze, & Purk, 1994; Lal & Rao, 1997; Shankar & Bolton, 2004; Shankar & Krishnamurthi, 1996).
- Second, under a High-Low pricing strategy, the retailer has to decide the advertising and communication effort applied to the initiative, understood as amount of investment from marketing budget allocated to inform consumers about the chosen price positioning.
- Finally and also related to High-Low strategies, the retailer has to determine the depth of the discounts to be offered, pricing positioning, which refers to the average discount magnitude applied to promotional items (Shankar & Bolton, 2004; Shankar & Krishnamurthi, 1996).

Some studies suggest that the pricing policy adopted by retailers is influenced by many diverse aspects (Voss & Seiders, 2003), among them factors related to the industry, the company itself, and other elements derived from the competitive situation and consumer demand.

When referring to a specific activity of sales promotions, such as price promotion, it is important to make reference to the DEC, which shows the representation of actual sales volume against price discounts applied during a certain period. Hence, the DEC shows pricing and volumes, and depicts pricing promotions effects over different products, such us private label and/or normal brands. Effects illustrated by the DEC can be basically grouped into three categories:

(1) The first category is related to direct discount effects. Two fundamental effects can be showed as far as this category is concern, namely, threshold and saturation.
Threshold stands for the minimum discount that has to be applied to ignite sales growth (Van Heerde et al., 2001), while saturation effect could be defined as the discount level that does not generate additional sales. This second effect can be justified either from the maximum number of product units that consumers can stock at home (especially with perishables products) (Blatterg et al., 1995), or from the consumer perception of discount itself, which has been shown to be lower than the real discount (Blair & Landon, 1981).
(2) A second category relates to the cross-effect generated from other products promotions. The cross-effects appear when other brands and categories promotion indirectly imply variation on the volume sold of a certain product. This variation could be different depending on the value assigned by consumer to the promoted brand (providing a much higher effect as the value perceived by the brand is higher)