Industrial Marketing Management 42 (2013) 544–551

Contents lists available at SciVerse ScienceDirect

Industrial Marketing Management

Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework

Jeroen D'Haen [1], Dirk Van den Poel *

Ghent University, Faculty of Economics and Business Administration, Tweekerkenstraat 2, B-9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 12 November 2011
Received in revised form 7 January 2013
Accepted 10 January 2013
Available online 29 March 2013

Keywords:
Customer acquisition
Sales funnel
Prospect
Nearest neighbor
Iterative framework

ABSTRACT

This article discusses a model designed to help sales representatives acquire customers in a business-to-business environment. Sales representatives are often overwhelmed by available information, so they use arbitrary rules to select leads to pursue. The goal of the proposed model is to generate a high-quality list of prospects that are easier to convert into leads and ultimately customers in three phases: Phase 1 occurs when there is only information on the current customer base and uses the nearest neighbor method to obtain predictions. As soon as there is information on companies that did not become customers, phase 2 initiates, triggering a feedback loop to optimize and stabilize the model. This phase uses logistic regression, decision trees, and neural networks. Phase 3 combines phases 1 and 2 into a weighted list of prospects. Preliminary tests indicate the good quality of the model. The study makes two theoretical contributions: First, the authors offer a standardized version of the customer acquisition framework, and second, they point out the iterative aspects of this process.

1. Introduction

The phrase *customer relationship management* (CRM) is often used in contemporary marketing literature. Although it has been in use since the beginning of the 1990s, researchers have reached no consensus with regard to its definition (Buttle, 2009a; Ngai, 2005; Richards & Jones, 2008). Most definitions have, however, some core features in common; for example, CRM consistently deals with the acquisition and retention of customers and the maximization of long-term customer value (Jackson, 2005; Ngai, Xiu, & Chau, 2009). Prior literature also distinguishes four types of CRM: strategic, operational, analytical and collaborative (Buttle, 2009a). This paper focuses on analytical CRM, which involves mining customer-related data for strategic purposes (Ang & Buttle, 2006; Buttle, 2009a; Ngai et al., 2009), centered on the process of acquiring new customers, and how data mining techniques can facilitate this process.

Most CRM literature neglects customer acquisition in favor of other topics, such as retention (Sohnchen & Albers, 2010), because retention strategies are typically cheaper than acquisition strategies (Blattberg, Kim, Kim, & Neslin, 2008a; Wilson, 2006). However, as important as customer retention might have become, customer acquisition is and should be a crucial focus for companies and researchers for several reasons (Ang & Buttle, 2006; Buttle, 2009b; Kamakura et al., 2005). Startups and companies aiming to exploit new markets need new customers, because they lack existing customers. Even existing companies in a mature market will lose some customers and must replace them (Wilson, 2006). Acquiring new customers is a multistage process, in which only certain suspects (for a definition of the terms used herein, see Section 2) become actual customers, also referred to as the "sales funnel" (Cooper & Budd, 2007; Patterson, 2007; Yu & Cai, 2007). During this process, it is often difficult for sales representatives to cope with all available data (Yu & Cai, 2007). Monat (2011, p. 192) indicates that many companies face this issue:

> "Sales leads are the lifeblood of industrial companies, yet determining which leads are likely to convert to bookings is often based upon guesswork or intuition. This results in a waste of resources, inaccurate sales forecasts, and potential loss of sales. A quantitative model that may be used to predict which leads will convert, based on information inherent in the leads themselves, would be highly valuable."

In response, this article presents a quantitative model, designed to be used as a tool to assist sales representatives in customer acquisition—that is, a sales force automation tool. Moreover, it is designed to be implemented in a web application, giving it certain specific characteristics and advantages. First, it should be usable regardless of specific company characteristics such as size and industry. Whether for a large company in the automotive sector or a small company in the food sector, the model should render high-quality predictions. Second, it must be fully automated and run without the need for human interference.

* Corresponding author. Tel.: +32 9 264 89 80; fax: +32 9 264 42 79.
E-mail addresses: Jeroen.DHaen@UGent.be (J. D'Haen), dirk.vandenpoel@UGent.be (D. Van den Poel).
URL: http://www.crm.UGent.be (D. Van den Poel).
[1] Tel.: +32 9 264 98 30.

0019-8501/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.indmarman.2013.03.006

Third, it must be fast and inexpensive. Because it is a web application, users typically want results immediately.[2] When the algorithm is embedded into a web application, the cost to the user is limited. The user (i.e., a business-to-business [B2B] company) only needs to pay a membership fee to obtain access to the application and does not need to pay for the whole database of prospects, which can be expensive. Moreover, the company does not need in-house experts to analyze the data, as the algorithm performs this step and provides intuitive, ready-to-use output.

Sales representatives must sometimes make arbitrary decisions in selecting prospects from a list of suspects and further qualifying them into leads. Thus, time is lost pursuing bad prospects and leads, violating the famous "time is money" corporate mantra. A model with high predictive power in forecasting the right prospects to pursue can save a company time and, ipso facto, money. Research indicates that approximately 20% of a sales representative's time is spent selecting prospects (Trailer, 2006) and depicts prospecting as the most cumbersome part of the selling process (Moncrief & Marshall, 2005). Furthermore, making ineffective decisions in the customer acquisition process decreases the overall value of the company over time (Hansotia & Wang, 1997). The proposed algorithm is designed to make the decision-making process less arbitrary by providing model-based prospects.

Although the algorithm should work regardless of the company using it or the industry in which it is situated, note that the proposed sales force automation tool will work best in markets that are highly saturated, in which market penetration is strategically crucial. We expect the highest efficiency in markets in which the pool of potential customers is large. In those markets, the selection process is often costly and arbitrary, due to information overload. In contrast, in industries in which customers are large organizations, well-known, and few in number, the proposed algorithm will not provide a significant advantage, because the selection of prospects is limited (Long, Tellefsen, & Lichtenthal, 2007). The algorithm functions in a B2B environment and uses the current customer base of a company to predict prospects. It also contains a feedback loop that iteratively improves its overall predictive performance.

There is a limited amount of research on customer acquisition (Blattberg et al., 2008a). With this research, we aim to fill this void and also stimulate further research. The theoretical contributions are twofold. First, we offer a standardized version of the customer acquisition framework. Second, we point out the iterative aspect of this process, which has been neglected in research. The remainder of this article is structured as follows: We present a literature review on customer acquisition, then describe the different stages of our model. After we elaborate on the data, we report the results of the model and finally discuss the conclusions, implications, limitations, and further research suggestions.

2. Customer acquisition framework

The sales funnel conceptualization offers a way to describe the customer acquisition process, dividing it into different stages (Ang & Buttle, 2006; Coe, 2004a; Patterson, 2007; Yu & Cai, 2007). These divisions vary from study to study, as do the definitions they use to characterize each part. A main difference, however, is where the studies place a prospect and a lead in the sales process: some put the prospect before the lead (e.g., Coe, 2004a; Metzger, 2005), whereas others put the lead before the prospect (e.g., Gillin & Schwartzman, 2011; Patterson, 2007). For the sake of clarity and as a way of creating a standardized framework, we first describe our vision on the sales funnel and define each stage. The emphasis is not on where the different terms are placed but on their definitions.

[2] We ran the algorithm discussed herein on a 3.40 GHz Windows server containing 16 GB of RAM.

The darker portion of Fig. 1 illustrates the sales funnel. The beginning is a list of suspects. *Suspects* are all potential new customers available. In theory, they could include every other company in a B2B context, apart from the current customer base. In practice, they boil down to a limited list of companies (perhaps purchased from specialized vendors; Buttle, 2009b; Rygielski, Wang, & Yen, 2002; Wilson, 2006). The vast amounts of information in those lists tends to overwhelm B2B marketers (Wilson, 2003). As a result, marketers often make selections using a set of arbitrary rules. The outcome of this selection is the list of prospects. *Prospects* are suspects who meet certain predefined characteristics. The next step is to qualify these prospects. *Leads* are prospects that will be contacted, after they have been qualified as the most likely to respond. This qualification is often driven by gut feeling or self-claimed competence. Finally, leads who become clients of the company are *customers*.

However, current theories and models fail to acknowledge the iterative nature of these stages, which implies none of the different stages is static. Yet the dynamics of this process influence the process itself. First, if customer acquisition is successful, the customer base is altered as new customers get added to it. As a result, these new customers are excluded from the next iteration in the sales funnel. Second, knowledge from a previous iteration should be incorporated in consecutive iterations. The successes and failures in each stage fine-tune the overall process. Here, we focus on the interplay between prospects and leads. The created model alters on the basis of the conversion from prospect to lead, including learning from the new information generated in each iteration. Incorporating the iterative aspect will improve the quality of customer acquisition models.

The procedure we propose radically alters the shape of the sales funnel (the lighter portion of Fig. 1), forming an isosceles trapezoid. More prospects are selected, but they are of higher quality. As a result, a greater proportion will be converted into leads and ultimately customers. Furthermore, the algorithm integrates a feedback loop that, over time, further elevates the quality of the prospects. Note that Fig. 1 is an exaggerated representation; reality should be somewhere between the graphs, because sales representatives will most likely select a smaller proportion of leads due to time constraints. It is nearly impossible for companies to increase their number of sales calls, assuming sales representatives work close to capacity (Coe, 2004b). The only alternative is to improve the quality of these calls, which is what the proposed algorithm aims to do. It provides high-quality prospects that are easier to convert, as recommended by research showing that call productivity can be improved by the use of information technology tools (Ahearne, Hughes, & Schillewaert, 2007; Eggert & Serdaroglu, 2011).

Traditionally, the conversion rate from prospects to qualified leads is approximately 10% on average (Coe, 2004b). Thus, getting a good list of prospects saves time that then can be spent qualifying them. Moreover, better qualified leads should lead to a higher customer conversion rate. Usually, a conversion rate from prospects to customers of 1%–5% on average can be expected (Coe, 2004b). Research shows that a lower conversion rate increases the cost of customer

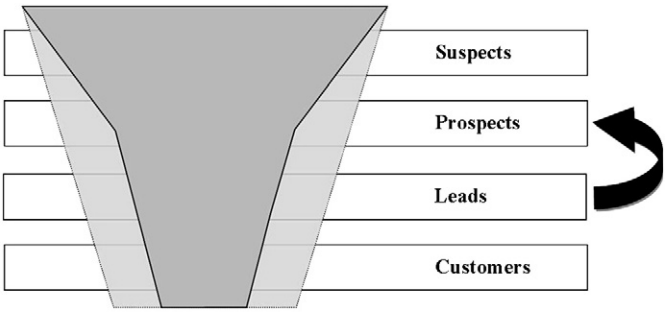

Fig. 1. The original and transformed sales funnel.