

Propuesta investigación Lead Scoring

idealista

Introduction

Personal selling is one of the most time-consuming processes performed by sales representatives (Sabnis et al. 2013). The process of developing prospects into customers is called lead management (Bradford, Johnston, y Bellenger 2016). The term lead is frequently used in the Marketing and sales departments, and can be defined as a contact generated by an existing or potential new customer that express interest in the company's offering or inquire into products or services (Todor 2016).

The prospecting process consists of a search for leads that identify potential customers, followed by the ranking or grouping of the identified leads in the order in which they are perceived to be convertible into orders. Based on the judgment or intuition of the salespeople, qualified prospects are people or organizations with needs or desires that can be satisfied by the seller's products or services. But prospects must exhibit a willingness or readiness to buy or at least a willingness to be exposed to the seller's offer. People with needs cannot be considered to be highly qualified prospects if they are resistant to sales calls. A third dimension is the prospect's ability to buy. This consists of authority to buy and buying power in the form of cash or credit. Yet, not all highly qualified prospects are described by a simultaneous desire for the seller's product, receptiveness to a sales call, and buying power. (Jolson 1988)

Targeting prospects is known to be quite unproductive since many leads reach low-probability prospects (Poppel 1983). Sales reps often argue that many leads lack the potential to result in actual sales, so they focus only on what they perceive to be "good leads" (Oliva 2006). Studies show that sales representatives never contact approximately 70% of the leads generated by marketing activities (Marcus 2002; Michiels 2009) resulting in what has been coined as a "sales lead black hole" (Hasselwander 2006). However, Hasselwander (2006) asserts that firms often lose "ready-to-buy" customers that sales reps never contact, and Moreau (2006) reports that disagreements about classifying leads can cause sales reps to discard customers that would be highly likely to buy in the future.

The overall result is a significant opportunity cost associated with poor lead follow-up, as well as demands for a viable solution to the "sales lead black hole".

According to Jolson (1988) prospects can be classified according to an arbitrary continuum of lead classification ranging from poorly qualified (loose leads in salespeople jargon) to highly qualified (tight leads) as Desire/Needs nonexistent, Desire/Needs unknown to the seller, Desire or needs are latent, Comparing alternatives, Preferred alternative and Ready to Buy. Typically, poorly qualified leads are quite costly in terms of lead cost per sale. This figure is calculated by multiplying the cost per lead by the average number of leads required per sale.

Leads at idealista

idealista is a real estate marketplace operating in Spain, Italy, and Portugal. The majority of ads at idealista have been published by professional real estate agents. When a seller wants to put their property up for sale, they'll generally contact a real estate agent to act on their behalf. With the seller's best interests in mind, the real estate agent will use their knowledge of the market to accurately price the property, list it and market it to potential buyers.

As an online marketplace, idealista sells slots to agents so that they can publish all the properties they are working on. Once the properties are in the listing potential buyers and tenants may contact the publisher to arrange and schedule property showings. These contacts, which may be written messages or phone calls, are considered leads. In many cases, real estate agents have no data to qualify the leads unless the seeker has included some relevant information in the body of the sent message. Once the sales team receives leads from idealista, they usually reach out to the individuals and determine if they meet the minimum criteria for becoming a sales opportunity.

Dealing with these leads is a time-consuming process for the agent and it may become especially hectic in the case of properties in high-demand locations where many leads per day are generated. A major expense for sales teams is the time wasted on dealing with a large volume of low-quality leads that will not become sales-qualified (Duncan y Elkan 2015). The number of marketing leads provided to sales reps can be so large that the sales reps cannot pursue all of them, even if they desired to do so.

Lead scoring is not new. Many companies currently use a manual lead scoring system (Duncan y Elkan 2015). Such methods are generally used by the marketing team to identify Marketing-qualified leads (MQLs). Marketing automation software facilitates the creation of such scoring systems and it has been around for more than 25 years (Moriarity 1989). In our case, idealista does not currently provide an associated score related for each lead generated to publishers. However, it is known that real estate agents rely on manual lead scoring systems, scores are hand-tuned by experienced members of the marketing or sales team. These systems typically use a "scorecard" in which the presence or absence of certain positive or negative customer attributes or behaviors are assigned fixed positive or negative values. These individual values are then summed to determine a final score for the lead.

Research question and goals

The purpose of this research is to calculate a score value that indicates how likely it is that a lead will result in a sale opportunity for the agent. The research question could be narrow as if idealista is able to build an automated lead scoring system. There is a great motivation to study this theme, specially because the lead score will allow real estate agents to discriminate between tight and loose leads. As a company, idealista is committed to providing the best experience to their customers. Salespeople generally prefer sales leads from prospects who are receptive ([Jolson 1988](#)), so we firmly believe that a good lead scoring mechanism would increase the satisfaction of those real estate agents who are customers of idealista.

The proposed research question is related to three main goals: - Identify relevant trends and insights from past converted leads - Identify which attributes are relevant to predict lead conversion at an early stage of the sales process - Develop an automated lead scoring algorithm that allows the real estate agents using idealista to adequately perform contact prioritization.

Identify relevant trends and insights from past converted leads

Identifying relevant insights from converted leads can be a challenging goal in the context of idealista. The actual process of selling or renting a property is not done inside idealista's ecosystem. When a particular real estate agent closes a deal with a property seeker they are not required to inform idealista about this successful operation. Usually, the publisher of the property in idealista removes the ad from the listing to avoid receiving more leads.

In this sense, for analytic purposes it can be assumed that a deal has been closed once an ad has been removed from the listing.

Identify which attributes are relevant to predict lead conversion at an early stage of the sales process

There might be different attributes relevant to predict lead conversion and they might be classified as static and behavioral features.

The static features are information about the property in the listing and the location where it is placed. The attributes of the listed ad include asking price information, the size of the asset, and other amenities. In addition, other location features can be directly retrieved from the analytics database. This information might include the expected time on the market of similar properties in the location and other information such as the average price in the area. These features can be directly retrieved from the ad that generated the contact.

Behavioral features represent actions taken by an individual and capture interest in the property by the potential customer during a specific period. These features might be numerical counts representing the number of times a user has performed a specific action that is tracked

by the portal. Examples of actions inside idealista might include visiting an ad page, favoriting the ad, forwarding them to seekers' friends/family, or requesting further information about the estimated value of the property.

Develop an automated lead scoring algorithm that allows the real estate agents using idealista to adequately perform contact prioritization.

Conventional and manual lead scores have some major drawbacks. They fail to capture nonlinear effects. For example, if a user contacts many properties, they will receive a high lead score, since they accumulate 5 points for each property contacted. It may even be the case that contacting many ads is a negative signal. For example, it could indicate the behavior of a person just grasping information on real estate data, or even a competitor, who is trying to attract the property owner to its services. In addition, complex interactions of features cannot be represented by scorecard models.

Another issue with conventional lead scoring is that the hand selection of values is error-prone. In particular, hand-selection is vulnerable to bias from potentially mistaken business logic. A third disadvantage is that traditional lead scores are unbounded positive or negative values. They do not intuitively map to the probability of lead conversion or opportunity close.

Automated lead scoring might be as simple as trying to reproduce existing linear qualifying criteria or using a Machine Learning model that learns beyond these rules. In any case, Duncan y Elkan (2015) proposes that the design of Automated Lead scoring systems must fulfill the following requirements:

- A model should be probabilistic and have a meaningful interpretation, such as expected revenue or probability of successful close.
- A model should not simply relearn an existing conventional lead classification
- A model should be consistent with a separate opportunity won/lost classification model. That is, it should assign higher scores to leads corresponding to closed-won opportunities than to leads that are not successfully closed.
- The model should be able to find quality leads quickly.

Once we have built the potential features of the model we could start training machine learning models. It is important to note that this is not a linear process. It is highly likely that after reviewing the results of the models we realize that we have to include further information to improve the models' performance, thus needing to build new features.

We also aim to Track the performance of the lead-scoring process by collecting the feedback provided by real estate agents for each lead-scored algorithm. The collected information would be useful to assess the effectiveness of the lead scoring system. It is important to note that experimentation with lead-scoring programs has reached different conclusions. In an experiment held in 1987, Jolson (1988) concluded that loose leads produced more sales volume, higher earnings for salespeople, a larger lead return rate, a lower lead cost per sale,

and a lower compensation-cost ratio. Tight leads produced a higher conversion rate into sales and an increased level of motivation for sales force members.

References

- Bradford, William R., Wesley James Johnston, y Danny Bellenger. 2016. «The Impact of Sales Effort on Lead Conversion Cycle Time in a Business-to-Business Opportunity Pipeline». *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2866954>.
- Duncan, Brendan, y Charles Elkan. 2015. «Probabilistic modeling of a sales funnel to prioritize leads». En *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015-August:1751-58. Association for Computing Machinery. <https://doi.org/10.1145/2783258.2788578>.
- Hasselwander, A. 2006. «B2B marketing confidential: B2B pipeline management».
- Jolson, Marvin A. 1988. «Qualifying sales leads: The tight and loose approaches». *Industrial Marketing Management* 17 (agosto): 189-96. [https://doi.org/10.1016/0019-8501\(88\)90002-8](https://doi.org/10.1016/0019-8501(88)90002-8).
- Marcus, Claudio. 2002. «Re-engineering lead management». *Gartner* 14 (3).
- Michiels, Ian. 2009. «Lead lifecycle management: Building a pipeline that never leaks». Research Report), Aberdeen Group.
- Moreau, R. 2006. «Five Ways to Improve your Lead Management».
- Moriarty, Rowland T. 1989. «Automation to boost sales and marketing». *Harvard Business Review*, 100-108.
- Oliva, Ralph A. 2006. «The three key linkages: improving the connections between marketing and sales». *Journal of Business & Industrial Marketing*.
- Poppel, Harvey L. 1983. «Marketing productivity and information technology». *Information Industry Insights* 5: 6-8.
- Sabnis, Gaurav, Sharmila C. Chatterjee, Rajdeep Grewal, y Gary L. Lilien. 2013. «The Sales Lead Black Hole: On Sales Reps' Follow-Up of Marketing Leads». *Journal of Marketing* 77 (enero): 52-67. <https://doi.org/10.1509/jm.10.0047>.
- Todor, R. D. 2016. «Marketing automation». *Bulletin of the Transilvania University of Braşov Series V: Economic Sciences* •. Vol. 9. www.data.com.