Unveiling the features of successful eBay smartphone sellers

Ana Teresa Silva[a], Sérgio Moro[b,*], Paulo Rita[c], Paulo Cortez[d]

[a] ISCTE Business School, ISCTE – University Institute of Lisbon, Portugal
[b] Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisboa, Portugal, & ALGORITMI Research Centre, University of Minho, Portugal
[c] Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal, & NOVA Information Management School (NOVA IMS), Campus de Campolide, 1070-312 Lisbon, Portugal
[d] ALGORITMI Research Centre, University of Minho, Portugal

ARTICLE INFO

ABSTRACT

The present study adopts a data mining approach based on support vector machines (SVM) for modeling the number of sales of smartphone devices by eBay sellers. The data-based sensitivity analysis was adopted for extracting meaningful knowledge translated into the relevance of each input feature for the model. Such approach allowed unveiling that the number of items the seller also has on auctions, the price and the variety of products the seller offers are the three features that influence most the number of sales, in a total of almost 25%, surpassing the relevance of the features related to customers' feedback.

1. Introduction

With the advent of Web 2.0 and online shopping, an immensity of data is collected from myriad applications and devices. EBay is an excellent example of an online company boosting its way through the Web 2.0 era, being currently one of the largest online sales platforms, supplying online retailing services for any seller worldwide (Einav et al., 2014). Such a colossal player entails a large set of different means for users to contribute with feedback on the services provided and registered sellers. These feedback data plus other relevant data (e.g., data on the items being sold, and users' characteristics) is scattered throughout multiple sources, which inevitably asks for some form of further treatment that allows classification, discovery of patterns and trends or prediction of outcomes. Such treatment implies the usage of increasingly complex and combined statistical and machine learning tools as the size of datasets builds up (Amado et al., 2018). Nowadays, datasets may extend to several Exabytes, increasing the challenging task of transforming such loads of information into actionable knowledge using adequate methods (Canito et al., 2018).

Data mining is the process of discovering patterns of knowledge from raw data (Sharda et al., 2018). Its roots lie on statistics and data analysis, and have been greatly enhanced through machine learning techniques and methods. Data mining as an evolving process has been around for some time, but only since the 1990s, when the concept was coined, until today has it been gaining considerably more popularity and attention (Fayyad et al., 1996; Sharda et al., 2018). This is happening due to the large amounts of data, in what is known as big data, that are generated every second from several sources, such as sensors and devices (Pal and Saini, 2014) and also social media and smartphones' applications (Chen et al., 2012).

These themes are the stepping stones for this data mining study. Therefore, its goal is to generate the type of information that is able to leverage decision support through actionable knowledge. It might be of particular interest for online retail sellers, online marketplaces and marketing practitioners, who may use the insights provided by the analysis of how online features of sellers' influence sales. In fact, large online e-commerce websites represent the future of retailers (Clemes et al., 2014), and top players such as eBay, Amazon and Alibaba are among the most technologically innovative organizations worldwide (Liu and Lu, 2015). Therefore, research on improving customer service based on cutting edge technology can help cope with the challenges of tomorrow.

Traditional data mining projects are time-consuming as all the data is often manually extracted and with limited amount of resources, which usually leads to limitations in the scope of analysis. In this case, the research is narrowed to the extraction of knowledge in the form of features' relevance from sellers of smartphones on eBay, one of the largest e-tailers worldwide (Kornberger et al., 2017). The aim of this study is to provide insights about what it takes to be a successful eBay smartphones' seller by unveiling through data mining which seller features contribute the most to actual sales, i.e. which have the highest influence on the number of items sold. Previous literature has approached the subject mostly from customer and potential consumers' perspectives yet rarely from the sellers' point of view. As the number of

registered sellers on online platforms rises worldwide, it becomes crucial to understand what drives the success of sellers within the different dimensions that can influence their results (Wu et al., 2015; Kannan, 2017). Such knowledge can be valuable both from a seller's perspective as well as for managing online platforms (e.g., offering premium services to the most prospective sellers or improving feedback services and information supplied to the registered users).

There is research focused on modeling the choice between auctions and posted prices (Einav et al., 2018), on online businesses emerging in the form of eBay ventures (Gregg and Parthasarathy, 2017), consumer trust in online purchases (Oghazi et al., 2018), e-satisfaction and consumer spending (Nisar and Prabhakar, 2017), and eBay sellers' reputation (Greenstein-Messica and Rokach, 2018). Yet a stream of research that grasps onto the conspicuous and measurable characteristics of online sellers combined with product attributes in order to determine their impact on sales using data mining predictive techniques is still scarce in the literature. Therefore, the immediate purpose of this paper is to fill in that research gap. In addition, the contributions for the literature are the following:

- Extraction of online seller features from an online sales renowned platform, eBay;
- Evaluation of the smartphones' online market through the analysis of eBay sellers' features and their impact on performance.

The next section dives deeply into the theoretical background, which supports the relevance of the subject along with the data mining techniques in use, followed by a detailed description of the chosen methodology and approach. Then, the results are discussed and interpreted in order to extract adequate knowledge out of the data. Finally, the conclusions are drawn in the last section.

2. Theory

2.1. Online sales

Web 2.0 is defined as a "set of applications and technologies that enable user-generated content, such as online social networks, blogs, video and photo sharing sites, and wikis" (Laudon and Traver, 2016, page 71). It is considered a new stage of development of the web and it differs from the previous one by the drastic increase in information density, interactivity and level of customization. This new phase can be traced back to 2007, when the changes became evident. It is also relevant to draw attention to the associated shift from making online purchases to going shopping online (Hemp, 2006, page 1) as the online environment and virtual communities become vital elements in the consumer journey, in a phenomenon often named "social commerce" (Huang and Benyoucef, 2013). Thus, recommendations from other consumers, instead of friend/family advice are also becoming an increasingly important decision factor (Kotler and Keller, 2012, page 138). It is important to examine the e-tail environment since "electronic markets enable volumes and speeds that human middlemen could not accomplish" (Venkatesan et al., 2006). However, there is still plenty of research focused exclusively on brick and mortar retail context when compared with pure online players and bricks-and-clicks, which have been growing expressively in the last years (Grewal et al., 2010).

Looking from the consumers' perspective, Cheung et al. (2005) pointed out that the main determinants of online consumer behavior were related with consumer characteristics, environmental influences, product/service characteristics, medium characteristics and merchants and intermediates' characteristics, which would have a transversal impact through the online customer journey.

In early research about pricing it was often argued that the advent of the Internet would lead to heightened competition online, which would induce price reductions (Brynjolfsson et al., 2006). Nevertheless, other features of online markets were found to have more impact in the

buying decision process such as variety and convenience. Hence, the trade-off between breadth and depth could stand a chance of being solved (Grewal et al., 2010). The turn up of Web 2.0 tools that speed up information sharing and networking has definitely been affirmative in self-generation of content.

Moreover, recent studies have devoted efforts in finding influencing features on the prices of online sales. Kocas and Akkan (2016) evaluated how online feedback and rating from customers affected the prices of books from twenty-four categories sold through Amazon.com. Their work has proved that customer ratings should be accounted for increasing profitability. Cao et al. (2015) presented a study on dynamic pricing of online shopping by dividing customers in patient and in-patient potential buyers, providing evidence that the optimal pricing policy should limit dynamic pricing when facing customers with little patience. Sellers' reputation has proven to be an effective influencer of the pricing policy followed, with highly reputed sellers having advantages in pricing, as shown by Xu and Ye (2015) through an analysis of TaoBao sellers. However, the same study also emphasizes that literature on pricing related to reputation is scarce. A previous article published by Ye et al. (2013) has reached a similar finding by analyzing both TaoBao and eBay sellers. Both studies are conclusive in that sellers' features do affect pricing, influencing sales performance, with the latter adopting a regression model for studying three sellers' attributes: reputation score, number of positive reviews, and score for "item as described". However, this study did not consider further features from sellers that are available on eBay, such as the neutral and negative reviews. Furthermore, both studies analyzed online sellers in a pricing perspective, not accounting for the number of sales derived from sellers' features, a research gap that the present study attempts to fill.

Managing the variety of products sold, i.e., the assortment has always been an essential element of business development (Ramdas, 2003). In today's environment where high levels of demand together with increased want for a personalized offer have become frequent, finding the right balance between variety and the level of customization is often a challenge. High variety can be associated with increased variability and lead to errors in forecasting (Ramdas, 2003; Fisher, 1997). The adoption of niche versus mass strategies is another important aspect related with assortment management and it is inextricably linked with the level of variety and specialization of the products sold. It was discovered that, from the demand side, huge variety of inefficiently organized items can stagger consumers and hold back purchases due to forecasting errors and difficulty for consumers to find the products they are looking for (Brynjolfsson et al., 2006). If sellers choose marketing and assortment strategies that are not compatible in ensuring a smooth supply chain (Fisher, 1997), it can have a negative impact on consumer behavior and repurchase intention based on satisfaction (Yen et al., 2007) which will inevitably affect sales.

2.2. Smartphones

Technologies and telecommunications have become essential elements of everyday life and business. The need for increasingly fast and optimized devices has guaranteed a steady growth in the technological industry, although at due different regional paces (Kellerman, 2010). Mobile devices have also become one of the primary sources for online shopping (Pearce and Rice, 2013). In the UK, for example, mobile has already surpassed desktop by 44% (The Guardian, 2016). Such relevance can prove to be an effective driver for increasing sales of mobile devices (Bilgihan et al., 2016). Smartphones belong to this category since they are essentially "mobile phones with more advanced computing capabilities and connectivity than regular mobile phones" (Statista, 2016). Although they have been available in consumer markets since the 1990s, only became truly popular and mainstream when, in 2007, the iPhone's introduction by Apple transformed the industry, leading also to the first Android based smartphone being released to consumer markets in late 2008 (Lee et al., 2015).