Australian Journal of Business and Management Research
New South Wales Research Centre Australia *(NSWRCA)*

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

A NOVEL SALES PREDICTION APPROACH: Competing vs Supporting Networks

Kun Chen
South University of Science and Technology

Dejun Xie
South University of Science and Technology

Huaiqing Wang
South University of Science and Technology

ABSTRACT

Modelling and predicting online sales based on the publicly available data from the online user community are important. In order to improve the accuracy of existing predicting model, we propose a weighted related networks approach to account for the contributions from related products in the same brand (supporting) and category (competing), plus other factors including the claimed past purchases by the user members, the number of reviewers and the relative prices. Three scenarios of such an approach are implemented, compared and tested with a baseline model that is developed based on a traditional approach in literature. The results show that categorical relationship impacts the online sales more significantly in general.

Keywords: *sales prediction, sales network, brand and category, online marketing*

INTRODUCTION

Enabled by Web 2.0 technologies, products are not only displayed but also assessed online. On the social media platform, product features associated with consumer experiences are easily found and become important factors affecting user buy. So the study of predicting online sales by social media information has been a hot topic.

Among the previous online sales studies, the consumer views, or the so-called "word-of-mouth", have been considered as a distinct factor for sales prediction. However, the product WOM is mostly measured as single-product-oriented, ignoring the fact that mutual product relationships would affect measuring single WOM. Such relations reflected on web pages are links that lead to each other. For example, you are currently reviewing a "Samsung" cell phone, after some minutes, you may be on an "Iphone" page through some clinks. So our research question is how to build the product relationships and whether the product network has impacts on product WOM and sales prediction.

To deal with the problems, we firstly identify three product properties from market theories, namely, product brand, product category and product attributes. Product brand is the "name, term, design, symbol, or any other feature that identifies one seller's product distinct from those of other sellers."[1] According to Chaudhuri and Holbrook [1], brand trust increases purchase loyalty, which in turn, leads to greater market share or product sales. Specifically, brand-loyal consumers are willing to buy a particular brand product because they perceive some unique value in the brand that no alternative can provide [2]. In online shops or consumer forums, it's a common-sense approach that products are grouped by brands. So brand is a primary factor to affect product sales. Product category is a way to organize products by their types. From consumers' perspective, choosing products from one category is a noncomparable process [3]. That means, a favor on a particular product often means less favor on other products in the same category, in other words, products are in competitive relations. Product category is also a common feature to organize products on websites, and it thus is another important factor to predict product sales. Product attributes are characteristics of a product that are thought to appeal to customers. In online shops, a product is always labeled by a group of keywords as attributes or features. Such labels are extracted from product reviews or added by consumers manually, and become a valuable source of information for customers making product choices online [4]. Products with more similar attributes indicate a stronger relationship between them.

[1] American Marketing Association Dictionary.

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

Above three properties organize products in a network. In this research, we will build the sales prediction model based on the product network to investigate the interaction effects of product properties and product WOM.

SALES PREDICTION REVIEW

In the literature, many approaches have been proposed for online sales prediction in social media with consumer reviews. One popular approach is to collect the nontextual measures such as numerical rating for sales prediction [4]. Other approaches based on content analysis and text mining, including single-entity (product)-oriented analysis, and multiple-product comparison (competitive) analysis [5], have also been used for measuring consumer purchases and product sales. Although the second group approaches, especially the multiple-product-oriented analysis, begin mining the product relationships in consumer reviews, the natural existing market structure and product properties are not considered as predicting factors in sales models.

The influence of brand on sales is an extension of the brand-human relationship, where the brand-human relationship defines a two-way interactive social relationship allowing a customer can change his loyalty in brand A to another brand B, if the intimacy and attractiveness of B is strong enough to overcome the original bond between the customer and the current brand. There have been considerable recent literatures to account for the mechanism and types of brand-consumer relations. For instance, Shimp and Madden (1988) decribes such a relation in three dimensions: intimacy, desire, and promise. Blackston (1992) clarifies this relationship into six types: admiration, reverence, threatening, cooperation, service, and disrespect. Aaker (1997) proposes five types of brand-customer relations as innocent, exciting, satisfactory, decent, and strong. Aggarwal (2004), however, views and classifies the relationship into two basic forms: exchange relationship and sharing relationship. Despite taking diverse perspectives and holding different views on how to define the brand human relationship, researchers generally agree that quality of the relationship is of critical importance to the long term and sustainable sales of any product. Accordingly, there are extensive studies devoted to the studying the factors which impact the quality of such relationship. Several most prominent factors include loyalty to the brand, strongness of the user community, image of the producer in general. For example, Thorbjornsen et al. (2002) uses empirical studies to conclude that a positive correlation is evident between the quality of the brand-customer relationship and the loyalty level to the brand. Algesheimer et al. (2005) empirically finds that the strongness of user social networks of a brand product is a telling predictor of the quality of the brand-customer relationship. Chaudhuri and Holbrook (2001) finds a positive correlation between the brand quality and emotional dependence, re-purchasing pledge, and user trust. In summary of these studies, the effect of competitors on the brand-customer relationship quality is evident and the competitor’s growth in one brand can be attributed to the decreasing of the brand image of the alternative producer.

Product category also plays an important role in determining the online sales because it can result in substitutes and complements effects of the products in the same category with different brands. An economically rationale buyer would prefer to buy the same quantity of utility in product by spending minimum amount of money (Russel and Peterson, 2000). A particular group of customers have consistent needs of a particular category of products. For example, online game players continuously buy innovated game products. Hikers regularly buy equipments for outdoor activities. Schools and universities need to order educational products routinely. One main reason hindering consumers from switching from one brand to another is usually the habit and convenience accumulated in the existing brand and switching cost. Once a competing brand in the same category is attractive enough in term of utility and economic benefit, or the convenience of keeping using the current brand does not compensate the price increasing in it, consumers will choose to switch brand (Molina-Castillo, et al. 2011).

Majority of consumers do not abandon the whole category of products in the contemporary world. This is especially true for the “global consumer class” who are well-educated in general and holding the top quartile purchasing power, and familiar with the trendy culture in business and society. They can switch from one brand to another, but they don’t abandon the whole product category since many of the products are required by their lifestyle or professional needs, thus are part of their routine and regular expenditure. And because such consumers are well-informed and rational in spending, the competing brands in the same category of products are usually defying each other. In other words, the users of the alternative brands have formed a fixed customer social community that share the same value and goodwill of that brand and individual members need to overcome the invisible but perceivable social and cultural barrier between groups to switch from one brand to another. Consequently, brands provided by alternative companies need to compete with the current brand for sales.

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

All the products in the different categories with different brands make up a network, where the information carried from the existing network can be used to predict the sales in the coming time period. Networks topology has been used for market structure and social network analysis in broad areas. However, the literature on product network analysis, especially on sales prediction, is very limited. Netzer, O. et al [8] use text mining method to build an undirected network based on online user-generated content to analyze the market structures and competitive landscape insights. Zhu Zhang et al [7] use customer messages to build a directed network to provide extra variables about product competitive relations for sales prediction. The product relationships defined in these researches is by measuring if the names or other equivalent identifications of the two products are contained in a sentence for a valid consumer's review. However, the approach is seriously challenged when two products are highly related but the names of which are not explicitly mentioned in a single sentence.

In the research discussed in the main text, we propose a product network topology for analyzing the interaction effects of product properties and product WOM, and develop a prediction model based on them. We show that the model's predicting ability is improved compared with previous work. In particular, we construct competing and supporting networks based on brand trust and category. Statistical analysis shows that WOM calculated based on competing-based network is better correlated with sales. We believe our approach and results have several important contributions to the literature in the field. First, it proposes a novel approach in sales prediction by constructing a competing and supporting network of the focal product. It fills the vacancies in many previous literatures where these factors are either lacking or treated as isolated parameters. Another significance of our approach is the formulation of adjusted WOM based on product attributes and keywords in reviews. It taps into review text instead of relying purely on the rating, thus, the model is more powerful.

RESEARCH METHODOLOGY: Competing vs Supporting Networks

The basic idea of our approach is that the demand for one particular product is determined by not only the WOMs of its own, but also the WOMs of other related products. The impact of the VOMs of related products can be positive or negative depending on whether the relationships between the two products are supportive or denial to each other. In the current stage, only two types of product relationships are considered. If the two products under study, say product A and B, are manufactured under the same brand or by the same company, then the calibrated WOM of B should exert positive effect on the WOM, therefore the prospect sales of A. On the other hand, if product A and B are classified in the same category, but not in the same brand, then the calibrated WOM of B should have negative impact on the WOM.

To illustrate graphically, suppose we have N similar products, denoted as P1 , P2, …, PN, among which the first N1 products are produced by the supporting manufacturers of P1 , and the rest N2 = N -N1 products are produced by the competing manufacturers of P1 , where the product P1 is the focus of study for the moment. In the following Figure 1, an edge segment connects P1 with each Pi , 2<=i<=N. Each of the (N-1) edges is colored either green or red, where green denotes a supporting relationship between the products the edge connects, and red denotes a competing relationship between the two products at the two ends. The scale of per unit WOM effect of Pi on the WOM of P1 is defined by the thickness of the edge C_{1i}(P1,Pi) ------the thicker the edge the more powerful the effect. The overall WOM effect of related products is the weighted average effect from all products P2, …PN , where the weight for the contribution from Pi, is calculated as percentage of attribute congruence between the product Pi and P1 for 2<=i<=N.

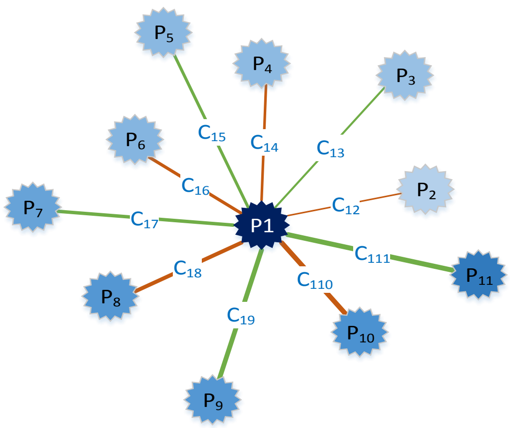

Figure 1. A graph model of product relationships

Australian Journal of Business and Management Research
New South Wales Research Centre Australia *(NSWRCA)*
Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

Assume P_1 earns M_1 reviews, each attached with a set of keywords or attributes, then the attribute set for P_1 is

$$\mathrm{Attr}(P_1) = \mathrm{Attr}_1^1 \cup \ldots \cup \mathrm{Attr}_j^1 \cup \ldots \cup \mathrm{Attr}_{M_1}^1$$

and then the attribute set for P_i is

$$\mathrm{Attr}(P_i) = \mathrm{Attr}_1^i \cup \ldots \cup \mathrm{Attr}_j^i \cup \ldots \cup \mathrm{Attr}_{M_i}^i$$

where $\mathrm{Attr}_j^1 = \{E_1^{1,j}, E_2^{1,j}, \ldots, E_k^{1,j}, \ldots, E_{L_j^1}^{1,j}\}$ is the set of keywords for P_1. The size of of set is L_j^1 corresponding to the j-th review to the product P_1. $\mathrm{Attr}_j^i = \{E_1^{i,j}, E_2^{i,j}, \ldots, E_k^{i,j}, \ldots, E_{L_j^1}^{i,j}\}$ is the set of keywords of P_i. The set size is L_j^i corresponding to the j-th review to the product P_i. Equivalently, $\mathrm{Attr}(P_1)$ and $\mathrm{Attr}(P_i)$ can be rewritten as

$$\mathrm{Attr}(P_1) = \{e_1^1, e_2^1, \ldots, e_j^1, \ldots, e_{L(P_1)}^1\}$$

$$\mathrm{Attr}(P_i) = \{e_1^i, e_2^i, \ldots, e_j^i, \ldots, e_{L(P_i)}^i\}$$

where $L(P_1) = \sum_{j=1}^{M_1} L_j^1$ and $L(P_i) = \sum_{j=1}^{M_i} L_j^i$. Next, we let $REL(P_1, P_i)$ be the relatedness between P_1 and P_i which is defined as the ratio between $\sum_{j_1=1}^{L(P_1)} \sum_{j_2=1}^{L(P_i)} 1_{(e_{j_1}^1 = e_{j_2}^i)}$ and the aggregate of $L(P_1)$ and (P_i). Then $REL(P_1, P_i)$ is key index to appraise the supporting and competing between P_1 and P_i[2]. In this way, the contribution of P_i to P_1 in WOM is calculated as

$$CON(P_1, P_i) = REL(P_1, P_i) WOM(P_i).$$

The total WOM contribution from all (N-1) products $P_2, \ldots P_N$ to P_1 can be calculated as the following averages.

(i) Simple average

$$RelatedCON = \frac{1}{N-1} \sum_{i=2}^{N} REL(P_1, P_i) WOM(P_i)$$

(ii) Weighted average

$$RelatedCON = \sum_{i=2}^{N} \frac{REL(P_1, P_i)}{\sum_{i=2}^{N} REL(P_1, P_i)} WOM(P_1, P_i)$$

Finally, the WOM variable value at P1 is computed as the summation of the WOM of its own and the weighted average WOM contribution from related products as computed from the above formula:

$$WOM\ (1) = WOM(P_1) + \frac{1}{N-1} \sum_{i=2}^{N} REL(P_1, P_i) WOM(P_i)$$

or

$$WOM\ (1) = WOM(P_1) + \sum_{i=2}^{N} \frac{Rel(P_i)}{\sum_{i=2}^{N} Rel(P_i)} WOM(P_1, P_i)$$

[2] We also use Jaccard and Cosine values in experiment. The results are almost same with the defined formula.

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

EMPIRICAL STUDY

The data source for the current study is Urcosme, an online community where users share their purchasing experience, feedbacks, comments, and selection tips on cosmetic products etc. It provides aggregated shopping intelligence to potential consumers based on product properties and customer reviews, including price, date of entry, category, brands, attributes, reviews, and available selling stores of cosmetic products. The source data is structured as interrelated multi-page networks, where the key parts for the current research are the consumer review page and product's profile page. Totally, we collect 451 product data on 177 brands and 16 category. A graphic model on 157 products is shown in figure 2.

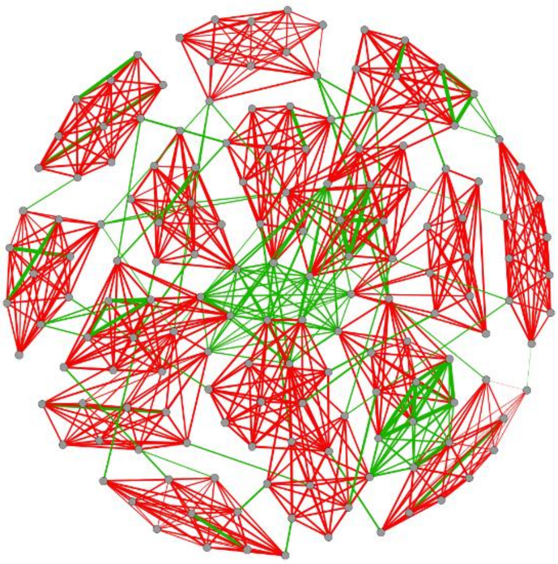

Figure 2. Cosmetics Product Network (157 nodes)

We choose potential buyers as dependent variable, because it measures the future sales performance. Supported by literatures such as [6, 8], the dependent variable is influenced by the price, previous sales record, number of reviews, and average rating (product WOM). A baseline model is thus built as:

Model 1

$LnPotentialBuyers = a_0 + a_1\text{LnPrice} + a_2\text{LnSalesRecord} + a_3\text{LnNumReview} + a_4\text{PRating} + \varepsilon$

Based on the proposed graphic model, we add the brand effect (positive relationship) to measure the product WOM, and build another model as:

Model 2

$LnPotentialBuyers = a_0 + a_1\text{LnPrice} + a_2\text{LnSalesRecord} + a_3\text{LnNumReview} + a_4\text{BrandRating} + \varepsilon$

Alternatively, if considering only the category effect (negative relationship) to measure the product WOM, the third model is built as:

Model 3

$LnPotentialBuyers = a_0 + a_1\text{LnPrice} + a_2\text{LnSalesRecord} + a_3\text{LnNumReview} + a_4\text{CategoryRating} + \varepsilon$

Finally, when incorporating both brand and category effects in the model, the model is written as:

Model 4

$LnPotentialBuyers = a_0 + a_1\text{LnPrice} + a_2\text{LnSalesRecord} + a_3\text{LnNumReview} + a_4\text{OverallRating} + \varepsilon$

ISSN: 1839 - 0846

Table 1 Results of comparison between models

Variable\Model	Model 1	Model 2		Model 3		Model 4	
		Simple	Weighted	Simple	Weighted	Simple	Weighted
Intercept	-4.470**	-2.665**	-1.938**	-3.801**	-1.041*	-3.585**	-2.128**
Ln price	0.137*	0.148*	0.159*	0.140*	0.111	0.139*	0.153*
Ln previousSales	0.260**	0.253**	0.239**	0.210**	0.245**	0.246**	0.239**
Ln NumReview	0.647**	0.689**	0.730**	0.710**	0.670**	0.667**	0.712**
PRating	0.619**						
BrandRating		0.194*	0.039				
CategoryRating				0.713**	0.762**		
OverallRating						0.465**	0.139*
Adjusted R^2	0.729	0.712	0.706	0.735	0.740	0.724	0.711
Significance Level: *p<0.01 **p<0.001							

The effectiveness of our model is evidenced by the above table, where the estimation of parameters and significance of the variables in p values, are provided. According to our empirical test, the potential sales can be properly explained by the past purchases disclosed by the UrCosme members, the number of sharings in form of online reviews, the selling price, and the impact from related products. The explaining power of the model, with an adjusted R^2 value of 0.729, is robust.

In order to account the impact from related products, three scenarios are considered to contrast the contributions from related products in the same brand, in the same category, and the inclusion of these two. Compared among the four models, we find the category relationship has a high correlation efficiency with the dependent variable, and the model build on such product relationship reaches the best performance (R^2 is 0.735 and 0.740 respectively). Another interesting and rather surprising observation of our tests is the band effect impairs the model performance. This conclusion is quite different from previous studies on brand trust, e.g. [1]. One possible reason is that the brand trust and effect mostly happens on individual buyers. When studying the product sales, it is not reasonable to consider brand effect on customers in general, but user characteristics and preferences should be included.

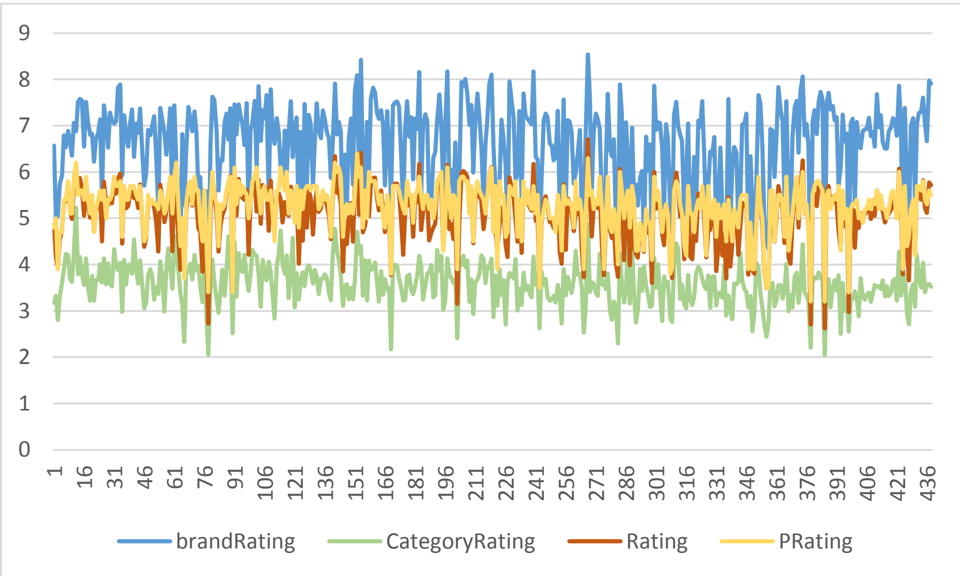

Figure 3. Comparisons between different product WOMs

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

When comparing the four product WOM metrics in Figure 3, it is clearly shown that the overall product rating is a combination of positive effects of brand and negative effects of category. According to model 3, the current product rating is a little higher than its actual value, because it missing the categorical competitive relationship between products.

CONCLUSION AND DISCUSSION

We present a novelty approach, through the current work, to explain and predict the online sales of cosmetics products. The model incorporates several key variables, including the claimed past purchase by forum members, the updated scaled selling price, the number of user sharings, and the impact from related products, where a related products networks approach is purposed and carried out throughout the analysis. The benefits of our approach, in comparison to those of the previously reported methods, have been sufficiently established by our empirical test. Nevertheless there are several points deserving further discussion.

The first aspect can be improved is to extend our investigation from single time step to multi-time steps. Put in another word, the analysis done so far has relied on a snapshot of the UrCosme web data to generate the projected sales in the next period. Since the website of UrCosme does not provide historical data, this extension will need us to fetch and record data periodically to accumulate a useful and large sample. But it will be a worthy try to test the merit of the model by time series data.

Another aspect to dig deeply is to sift all the online opinions and sharings. This will better reflect the market reality since all the community members or users are very different from personal preference, knowledge and experience of beauty, willingness to share information, buy habits, price sensitivity, etc. In some sense, the connectedness and weights analysis for review writers can be combined with product relatedness for WOM based online sales.

Finally, other potential applications on product relationship networks include mining customer behaviors, discovering market strategies and predicting product popularity.

REFERENCE

1. Chaudhuri,A., Holbrook, M.B., “The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty”, Journal of Marketing: April 2001, Vol. 65, No. 2, 2001, pp. 81-93.
2. Pessemier, E.A., “A New Way to Determine Buying Decisions,” Journal of Marketing, 24, 1959, (October), 41-46
3. Johnson, M.D., “The Differential Processing of Product Category and Noncomparable Choice Alternatives”, Journal of Consumer Research, Vol. 16, No. 3, 1989, pp. 300-309
4. Archak, N., Ghose,A., Ipeirotis, P.G., “Deriving the Pricing Power of Product Features by Mining Consumer Reviews”, Management Science, 57, 8, 2011, pp. 1485-1509
5. Chevalier, J.A. and Mayzlin, A., “The effect of word of mouth on sales: Online book reviews”, Journal of Market Research, 43, 3, 2006, pp.345-354.
6. Zhang, Z., Li, X., and Chen, Y., “Deciphering word of mouth in social media: Text-based metrics of consumer reviews”, 2012, ACM Transactions on Management Information Systems, 3, 1.
7. Zhang,Z, Guo,C. and Goes, P, “Product comparison networks for competitive analysis of online word-of-mouth”, ACM transactions on Management Information Systems, 3, 4, 2013, pp.20:1-20:22
8. Netzer,O, Feldman,R Goldenberg,J, and Fresko,M. “Mine your own business: market-structure surveillance through text mining”, Market Science, 31, 3, 2012, pp. 521-543
9. Blackston T. (1992). “Obserations: building equity by managing the brand’s relationships.” Journal of Advertising Research, 32, 1992, pp. 101-105.
10. Blackston M. (1995). “The qualitative dimension of brand equality.” Journal of Advertising Research, 7, 1995, pp. 35-43.
11. Shimp T. and Madden T. “Consumer-object relations: a conceptual framework based analogously on Stemberg’s triangular theory of love.” Advances in Consumer Research, 15, 1, 1988, pp.163-68.
12. Aaker D. A. “Discovering brand magic: the hardness of the softer side of branding.” International Journal of Advertising. 16, 1997, pp. 199-210.
13. Thorbjornsen H., Breivik, E., and Supphellen, M. “Consumer-brand relationships: a test of alternative models.” Conference Proceedings of American marketing Association, 13, 2002, pp. 283-85.

Vol.4 No.10 | February-2015 ISSN: 1839 - 0846

14. Algesheimer, R. Dholakia, U., and Herrmann, A. "The social influence of brand community: evidence from European car clubs." Journal of Marketing, 69, 3, 2005, 19-34.

15. Cleeren, K., Dekimpe, M, and Helsen, K. "Weathering product-harm crises." Journal of the Academy of Marketing Science, 2008, 36, 2, pp. 262-70.
16. Shock, A. D., Bayus, B. L., and Kim, N. "Product complements and substitutes in the real world: the relevance of other products". Journal of Marketing. 68, 2004, pp. 28-40.
17. Molina-Castillo, F. J., Munuera-Aleman, J. L., and Calantone, R. J. "Prodcut quality and new product performance: the role of network externalities and switching cost". Journal of Product Innovation Management. 28, 6, 2011, pp. 915-29.
18. Keller, P. and Amstrong, G. "Principles of Marketing". 10th edition, Pearson Education