UNDERSTANDING THE DRIVERS OF TOURISTS EXPLICIT RECOMMENDATIONS

1. Introduction

One of the most challenging decisions of every tourist when arriving at an unknown destination is to choose the places to visit. Recommendation sites such as TripAdvisor or Yelp are crucial to guide tourists on where to find the places that best fit their expectations (Litvin, 2015). One such example is the choice for a restaurant and how to find a place to enjoy a good meal while in a foreign city. While exploring reviews, consumer look for how their peers evaluate a given experience and particularly if they explicitly recommend it to others. Therefore, exploring what drives tourists to write a review with an explicit recommendation may help managers to focus their operational efforts on a better customer satisfaction on such drivers.

Traditional methods use surveys to ask consumers what motivates them to issue a recommendation (Filieri, 2015). However, today, there is a plethora of opinions being written online that may bring new perspectives to the antecedents of consumers recommendations. Such unstructured data, although a valuable source of information, must be treated with a systematic method that may be able to automatically structure and treat millions of reviews into a model which may predict consumer recommendations.

The current paper uses a text mining approach to structure online reviews using a lexicon based on positive and negative markers. Such markers are then used to find those that significantly affect explicit recommendations (either positive or negative) in online reviews.

The paper begins with a literature review on the most important findings about consumers' motivations to write online reviews that may help others to find the best place to visit. The method used to gather and treat the unstructured data is then explained in the methodology section and the results section shows the main findings of the current study. Finally, results are discussed and conclusions highlight the study limitations and point to future research paths.

2. Literature Review

Today, online recommendations are one of the most important sources of information to help consumers in their decision-making process (Filieri, Alguezaui & McLeay 2015; Fotis, Buhalis & Rossides, 2012; Sparks, Perkins & Buckley, 2013). Consumers perceive opinions of their peers as having higher value than those posted by the provider (Bickart & Schindler, 2001). A good review often helps consumers decrease the risks associated with choosing a restaurant or an hotel that does not meet their expectations (Bronner & de Hoog, 2011). The reviewers that gather the higher number of fans are usually those who alert their followers about things to avoid in a given experience or that highlight the most pleasurable items to look for (Guerreiro & Moro, 2017).

Motivations to write consumer reviews stem from personal interests, social benefits, social concern, functional purposes, quality assurance motivations, economic incentives, entertainment and the motivation to help the company for whom the review is written (Bronner & de Hoog, 2011). Being helpful to others by sharing positive experiences (social concern) or helping peers to avoid problems that may arise while making their consumption decisions and thus helping them to save precious time (functional) are often mentioned in the literature. Such advices all contribute to consumer empowerment while they face the decision to choose the best options available, especially while traveling abroad.

Online reviews are generally free of mandatory topics. Reviewers may choose to write what they think better expresses what they have experienced. Consumer's opinions may have a more negative or a more positive tone when they discuss service attributes such as their feeling about the experience, the attitude of the staff or how much time they have to wait in line to get served, just to name some examples. However, only some of the reviews explicitly recommend the service to their peers. Sentences such as "I highly recommend this restaurant" or "I will certainly come back" are sometimes included in the reviews to show peers that the service is a worthy experience. Understanding what drives reviewers to explicit recommend a service may help managers to selectively improve their offerings to better meet consumer's expectations. However, due to the unstructured characteristics of online reviews, a model that may predict the antecedents of explicit recommendations in text is often difficult to obtain without using some kind of formal steps to structure such information. Text Mining has been used successfully in recent studies to overcome such challenges (He, Zha & Li, 2013; Guerreiro & Moro, 2017).

Text Mining is a semi-automated process of transforming unstructured text into a structured form of data that may later be analysed in search for hidden patterns (Mostafa, 2013). Unlike structured data, in which all variables are well determined and their value is a continuous or categorical value, in text, all information may lead to relevant pieces of knowledge that was not previously determined.

Today, most information online, particularly valuable information such as consumers opinions are in textual format. Due to the huge amount of text being written every second online, reading and interpreting all those opinions would be impossible. Without a formal set of rules to automatically treat such information, results may easily become subjective.

Text mining presents a set of formal steps that help researchers in structuring textual information and has been used successfully to discover patterns in text in many wide fields of study such as in literature review analysis (Guerreiro et al., 2016), term extraction (Zhang, 2008), and opinion analysis of online reviews for the tourism industry (Pekar & Ou, 2016).

A first step in text mining, usually involves treating text to be sure that words that have the same meaning are treated as the same word in the analysis. Such treatment method is called stemming and extracts the radical of each term so that words such as *tourist* and *tourists* are treated as the same (Porter, 1980). A natural language processing (NLP) algorithm called part-of-speech is often applied to keep the context of semantical information of each sentence, after which, auxiliary terms, called stopwords that are not relevant to the analysis are removed. In the end, the unstructured text may be represented in a document-to-term matrix (DTM) that crosses relevant terms with their frequency in each document (in the current case, in each review).

Text mining also includes steps to analyse the sentiment score of textual elements. Although it's important to understand what consumers are discussing online, it's even more valuable to highlight if they are writing in positive or negative terms about a given subject. Therefore, sentiment analysis methods are often applied to classify words or sentences into negative, neutral or positive sentiments. The lexicon-based approach is a method used in sentiment analysis to classify such polarities. In a lexicon-based approach a set of dictionaries hold information regarding the strength of each word within a sentence (Liu & Zhang, 2012).

A final DTM may hold not only how frequent each term is in consumer discussions, but also their sentiment score.

Researchers are also frequently interested in measuring the impact of some issues discussed in text on some dependent variables, for example, in consumer rating. For such purpose, classification methods such as logistic regressions and decision trees are commonly used with large samples (Wu et al., 2008). However, despite the power of such methods in finding useful patterns in data, they present some challenges.

Although logistic regressions are well fit to model dichotomous dependent variables, they lack the ability to show the boundaries of the independent variables that better explain the changes in the dependent one. Another limitation of using logistic regressions with large samples such as the ones found in online reviews is that even minor effects may become statistically significant which may endanger the interpretation of results (Lin et al., 2013). Decision trees (DT), on the contrary, perform well with large datasets and a large number of input variables. The use of a splitting algorithm such as the information gain method or the Gini algorithm commonly used by decision trees, only selects the input variables that better separate the several classes of the output attribute (Rokach & Maimon, 2014). Therefore, only the variables that better explain the dependent variable are included in the model. Finally, decision trees are very easy to interpret and explain. Therefore, they are often used to generate rules to support managerial decisions. However, decision trees also have their disadvantages over logistic regressions. For example, DT assumes that decision boundaries are orthogonal to the axis, which limits the possibility to model very complex non-linear relationships. Also, the algorithms used to grow the DT are hierarchical, which means that they create nested hyper-rectangles to represent the feature space up until they reach a stopping rule. Such effect may lead to overfitting if too many branches are used to grow the tree (Brieman et al., 1984; Schaffer, 1991).

3. Method

This study used a dataset available from Yelp (Yelp, 2017) that contained customers' reviews from more than 144 thousand different businesses in more than 11 cities across 4 countries. A sample was extracted that only contained restaurants. A text mining approach was used to process unstructured text into structured textual data (Miller, 2005) and sentiment analysis based on natural language processing (NLP) was used to understand the context of each opinion. A group of dictionaries available in IBM SPSS Modeller Text Analytics, a commonly used tool for sentiment analysis (He et al. 2013), were used to find the reviews with an explicit recommendation on the review text and use it as the dependent variable. The recommendations dictionary available searched for words such as the example n-grams "come back" and "give them a try" for positive recommendations (classified as 1) and "can't recommend" and "do not plan to return", for negative recommendations (classified as -1). Independent variables were also derived from available dictionaries for opinion analysis. Reviews were classified as 1 (presence) or 0 (absence) of n-grams related to positive and negative attitudes, feelings, competences, functioning, budget, wait-time, and customer support. For example, all reviews where consumers complained about the wait time in the restaurant were classified using words such as "wait in line", "queue" while positive and negative words related to budget (e.g. cheap, affordable) and competence (e.g., able to resolve, efficient) were also binary classified according to those different markers.

The final dataset contained 132,657 reviews in which 71% had positive and 29% had negative explicit recommendations. A decision tree algorithm (CHAID) and a binary logistic regression were used to predict the drivers that would influence an explicit recommendation in the text. The data was divided into a training and a testing sets and balanced between positive and negative recommendations so that models were compared for their fit to predict tourist recommendations.

4. Results

The binary logistic regression using a stepwise method for variable selection, explained 21% of the variance in recommendations (Nagelkerke R^2) and correctly classified 68% in the training sample and 62% in the test sample.

Cox and Snell	.162
Nagelkerke	.216
McFadden	.128

Table 1. Logistic Regression Model Pseudo R-Square

Results show that positive feelings have the highest positive effect on recommendations (Exp(B)=2.975, p<.05) and negative competences have the highest opposite effect (Exp(B)=.293, p<.05).

Table 2 shows the logistic regression coefficients.

	B	Std. Error	Wald	Sig.	Exp(B)
Intercept	-.187	.017	122.192	.000	
Customer Support	-.210	.022	94.285	.000	.810
Negative Attitude	-.953	.034	806.632	.000	.386
Negative Budget	-.671	.032	448.383	.000	.511
Negative Competence	-1.227	.126	94.383	.000	.293
Negative Feeling	-.855	.026	1078.451	.000	.425
Negative Functioning	-.479	.026	343.949	.000	.620
Positive Attitude	.816	.026	1020.085	.000	2.261
Positive Budget	.116	.036	10.452	.001	1.122
Positive Competence	.212	.037	33.333	.000	1.236
Positive Feeling	1.090	.021	2769.146	.000	2.975
Positive Functioning	.376	.069	30.039	.000	1.456
Positive Store	-.064	.029	4.839	.028	.938
Positive WaitTime	-.244	.028	76.722	.000	.784

Table 2. Logistic Regression Coefficients

A CHAID decision tree using a Bonferroni adjustment (67% reviews correctly predicted in the training data and 62% of reviews correctly predicted in the testing data) show that feelings and attitudes are the most important predictors for recommendations in a review. When customers write comments that show a negative attitude (e.g. arrogant, bad-tempered), 83% of them write a review that has an explicit negative recommendation. On contrary, having a review that has positive feelings (e.g., attractive, clean), and no negative feelings (e.g. bitter, dirty) or negative attitudes, leads to a positive recommendation in 72% of the times.

Figure 2 shows the complete decision tree using three levels below root as a stopping criteria.

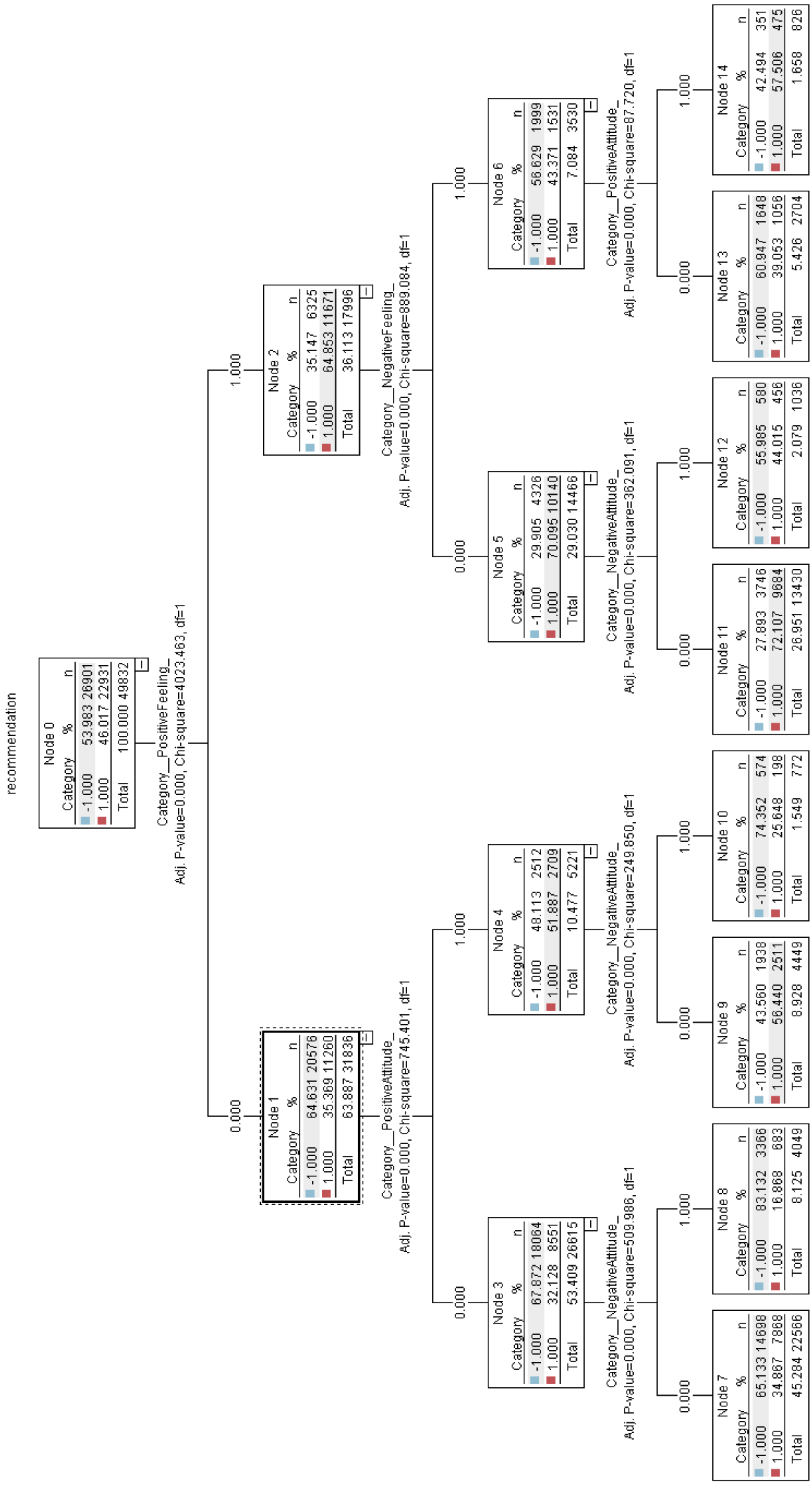

Figure 2. CHAID Decision Tree

5. Discussion and Conclusion

The current study shows that feelings, attitudes and competences may be important predictors for an explicit recommendation in the text of reviews, which may help managers to deal with such markers proactively due to their effect on one of the most important parts of WOM, the act of recommending.

The study shows that the lack of competences and negative attitudes are important drivers to predict when a reviewer issues a negative recommendation on the body of the review. For example, if consumers feel the staff had a bad attitude when they attended them, usually they recommend their peers to avoid the service. The same effect occurs if consumers felt that the provider was not able to efficiently resolve problems that occurred while they were being served. The findings align with studies from Stringam & Gerdes (2010). They found that words such as “apology” and “refund”, which suggest a lack of competence from the provider were negatively associated with consumer ratings. Our study suggests that such effect may also lead to a negative explicit recommendation in the review.

Zhang et al. (2010) show that cleanness and decoration are positively associated with the popularity of restaurants. Our research aligns with such findings and also shows that positive feelings usually predict a positive recommendation in the body of the review, particularly when there are no negative feelings to share as well. If consumers are completely satisfied with the service not only on functional terms, but especially if the service highlighted positive feelings such as attractive decoration or a clean environment (e.g.), they usually recommend it to their peers.

Despite the contributions of the current study, there are some limitations that must be acknowledged. First, although the sample size was high, the study used a single source of information (Yelp) and a single type of business (restaurants). Further research may confirm the current results using other sources and reviews about other types of services in hospitality and tourism. Second, groups of terms were derived using a lexicon of terms already embedded in SPSS Modeller grouping terms around recommendations, feelings, attitudes, competences, functioning, budget, wait-time, and customer support. Future work may use topic modelling to cluster terms automatically using the co-occurrence of words inside each review (Blei & Lafferty, 2007)

Although the study only explains part of the recommendations variance, it points to important markers for future research. One possible avenue for further analysis may be to use the sentiment score of feelings, attitudes and competences instead of only binary classifications.

REFERENCES

Bickart, B., & Schindler, R. M. (2001). Internet Forums as Influential Sources of Consumer Information. *Journal of Interactive Marketing, 15,* 31-40.

Blei, D. M., & Lafferty, J.D. (2007). A correlated topic model of science. *The Annals of Applied Statistics, 1*(1), 17-35.

Breiman, L., Friedman, J., Olsen, R., & Stone, C. (1984). *Classification and Regression Trees.* Wadsworth and Brooks.

Bronner, F. & de Hoog, R. (2011). Vacationers and eWOM: Who Posts, and Why, Where, and What? *Journal of Travel Research, 50*(1), 15-26.

Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travellers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management, 51,* 174-185.

Fotis, J., Buhalis, D., & Rossides, N. (2012). Social media use and impact during the holiday travel planning process. In M. Fuchs, F. Ricci, & L. Cantoni (Eds.) *Information and communication technologies in tourism 2012* (pp. 13 - 24). Vienna, Austria: Springer-Verlag.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management, 33*, 464–472.

Guerreiro, J., Rita, P., & Trigueiros, D. (2016). A Text Mining-Based Review of Cause-Related Marketing Literature. *Journal of Business Ethics. 139*(1), 111-128.

Guerreiro, J., & Moro. S. (2017). Are Yelp's Tips helpful in building influential consumers? *Tourism Management Perspectives. 24*, 151-154.

Lin, M., Lucas, H. C., & Shmueli, G. (2013). Too big to fail: Large samples and the p-value problem. *Information Systems Research, 24(*4), 906-917.

Litvin, S. W., Blose, J. E., & Laird, S. T. (2005). Tourists' use of restaurant webpages: Is the internet a critical marketing tool? *Journal of Vacation Marketing, 11*(2), 155–161.

Liu, B., & Zhang, L. (2012). *A survey of opinion mining and sentiment analysis. Mining Text Data* (pp. 415–460). Springer US.

Miller, T. W. (2004). *Data and Text Mining: A Business Application Approach.* New Jersey, NJ: Prentice-Hall, Inc.

Mostafa, M. M. (2013). More than Words: Social Networks' Text Mining for Consumer Brand Sentiments. *Expert Systems with Applications, 40*(10), 4241–51.

Pekar, V., & Ou, S. (2008). Discovery of subjective evaluations of product features in hotel reviews. *Journal of Vacation Marketing, 14*(2), 145-155.

Porter, M. F. (1980). An Algorithm for Suffix Stripping. Program. *Electronic*

Library and Information Systems 14(3), 130–37.

Rokach, L., & Maimon O. (2014). *Data mining with decision trees: theory and applications*. World Scientific Publishing.

Schaffer, C. (1991). When does overfitting decrease prediction accuracy in induced decision trees and rule sets? *Machine Learning—EWSL-91.* Springer Berlin/Heidelberg.

Sparks, B.A., Perkins, H.E., & Buckley, R. (2013). Online travel reviews as persuasive communication: the effects of content type, source, and certification logos on consumer behaviour. *Tourism Management, 39*, 1-9.

Stringam, B., & Gerdes J. (2010). An Analysis of Word-of-Mouse Ratings and Guest Comments of Online Hotel Distribution Sites. *Journal of Hospitality, Marketing & Management, 19*(7), 773-796.

Zhang, W., Yoshida, T., Ho, T. B. & Tang, X. (2009). Augmented mutual information for multi-word extraction. International Journal of Innovative Computing, *Information and Control, 5*(2), 543–554.

Zhang, Z., Ye, Q., Law, R., & Li, Y. (2010). The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management, 29*, 694-700.

Yelp. (2017). Yelp Challenge Dataset 2017. Retrieved from https://www.yelp.com/dataset_challenge

Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., Steinberg, D. (2008). *Top 10 algorithms in data mining*. *Knowledge and Information Systems*, *14*(1), 1-37.