COMPARATIVE STUDY OF OPINION MINING IN A RETAILER NETWORK

Andrea Hecksher[1], Nelson F. F. Ebecken[1]

1 Rio de Janeiro Federal University (UFRJ)

ABSTRACT

This present article compare results acquired by an analysis performed based on Text Mining techniques of contents extracted from the *fan page* of a retailer store network, comparing with results found in field research done for the same retailer network. The research intends to corroborate to the idea that Text Mining techniques, in special, the Feelings Analysis, which present themselves everyday more as an effective and advantageous solutions for the understanding of opinions, studies of perception, and brand reputation.

Keywords: Opinion mining, Retailer Network, Feelings Analysis, Client Satisfaction.

1. INTRODUCTION

Analyzing the reputation of brands, capturing the opinion of consumers, and acquiring the understanding regarding their preferences related to products and services are the main part of the work of marketing researchers and professionals. However, traditional market research methods, such as focal groups and face-to-face interviews, are expensive and time-consuming.

It is already a reality for any brand manager that the digital media and social networks are important for branding[1] strategies and for company advertising campaigns. Nowadays, large companies use social media to interact with clients, aiming to improve sales, consumer loyalty and retention, building up a consciousness and reputation around the brand. By generating informative or promotional content, enterprises aim does increase, retain, and build a loyal relationship with the target audience. The use of social media enables activities such as: brand management (through publicity, marketing, and content distribution), sales, costumer care and support, product development and innovation (Culnan; Mchugh; Zubillaga, 2010; Di Gangi; Wasko; Hooker, 2010 *apud* He; Zha; Li, 2013).

1   "Branding is a hybrid subject, a result from the association of areas in marketing, advertisement, and design, and that deals with management, communication, and format, in that order" (GOMEZ, 2009 apud GOMEZ et al., 2012).

Everyday there is an increase of the number of companies that use social media, such as *Facebook* and *Twitter*, to provide services and to interact with clients. At the same time, there is an increasing number of clients that interact with companies and brands through these same channels, or that spread their opinion in different websites and social media applications. As a result, a large amount of content spontaneously generated by the users of social media is available for free. To effectively evaluate their business environment, companies must monitor this mass of information and transform the social media data into understanding for managers and marketing professionals.

The internet has facilitated the access to information and socialization. The free availability of search engines and websites with engines of price and product comparison has provided consumers with more information, making them more demanding. The reports from consumers and the information exchange under virtual communities elevate the content generated by users as a meaningful source to analyze interests and opinions of clients regarding the company, its products and services. As a whole, consuming opinions and experiences reported by consumers are more credible for other future clients than ad campaigns, or any other marketing campaign strategy. Virtual discussions have a considerably large influence in the perception of a brand or in choosing a product.

The smart use of social media by companies does not resume into having a promotional and communications

channel with their public. Besides that, there is a fundamental role to promote an always updated perception of the opinion of consumers regarding products or services. Then, blogs and social network websites can be understood as measurement tools and indicators of brand reputation. Through tracking and extraction of relevant information in posts and tweets, companies can immediately intervene in the communication with unsatisfied clients, using consumer comments as feedback, and identifying preferences and demands for services and products.

Based on the fact that around 20% of microblogs mention the name of a brand (Jansen *et al.*, 2009), the management of the image of the brand in social media must be an included part of the marketing strategy of a company. Digital social networks are the richest and free source to learn opinions and critics, from users and consumers, regarding products and services.

Based on the described scenario, the present exploratory research pretends to corroborate with the idea that Text Mining techniques, in special the Feelings Analysis, are presented more and more as effective and advantageous solutions to both know the opinions and studies of perception and reputation of brands, and to confirm the characteristics of practicality and accessibility of studies regarding time and costs. The conclusion emerges from a comparative analysis between the results acquired from a field research performed in a retail trade network located in the city of Rio de Janeiro, and the results found from an analysis done using Text Mining from contents found on the *Facebook* page of the mentioned network. The goals, both from the field research, from the mining, and from the company social network page, consisted in acquiring as much information as possible regarding the opinion of the clients about the stores, the customer service, and the mix of products offered by the company, in order to acknowledge the strong and weak points of the enterprise.

2. RESEARCH METHODS

To measure the perception of the clients of the retailer network, there was an exploratory research using Text Mining techniques, applied in contents removed from the the retailer network fan page.

For opinion mining and analysis of textual content, two Text Mining techniques were used: Keyword Extraction and Feelings Analysis.

2.1 Text Mining

Text Mining is an extension from Data Mining. Data Mining refers to the extraction of understandings from database that can have many different types of information setting, such as numerical, text, hyperlinks, and other multimedia formats. Text Mining, also known as Text Data Mining (Hearst, 1999 a*pud* Ranfagni; Guercini; Camiciottoli, 2014), concentrates on the extraction and elaboration of information strictly from textual sources (Witten, 2005 *apud* Ranfagni; Guercini; Camiciottoli, 2014).

Text Mining is the stage of the analysis of the process known as KDT (Knowledge Discovery in Textual Databases), and it can be interpreted as a process of textual contents that usually employs statistical techniques of computer algorithms to find a non-structured text data information. The stages of the process of Text Mining include: pre-processing, adaptation of Data Mining algorithms to be applied in Text Mining, gathering information from the internet, and observation of the information.

According to Tan (1999 *apud* He; Zha; Li, 2013), Text Mining has a high commercial potential, once texts are the most natural format of information storage and exchange. In 1999, Tan already mentioned a study which indicated that round 80% of the information from an organization was present in text documents, such as emails, memos, client mail, and reports. From that moment, the volume of text information only increases, and today, the largest challenge is the treatment of the enormous load of textual data. Therefore, to extract useful information from a large amount of text documents, it is imperative the use of computer automated techniques (He, 2013a; Liu; Cao; He, 2011 a*pud* He; Zha; Li, 2013).

Text Mining is applied mainly to clustering, text summaries to extract information, and the analysis of links. Clustering, or cluster analysis, in special, has an advantage to spot standards, tendencies, and correlations. Clustering is the non-supervised classification of data, generating groupings, or clusters. It represents one of the staged of data exploratory analysis, also called Cluster Analysis (Jain; Murty; Flynn, 1999).

Text Mining techniques have been applied successfully in researches and analysis in many areas, such as in business, health sciences, education, search and recovery of information as in library and information sciences. Despite that, in many cases, the application of Text Mining aims to extract new and relevant information in a large amount of non-structured texts.

Simple examples of Text Mining that are used more frequently in websites are the Word Clouds or Tag Clouds, in which they provide a visual summary of the topic contents of the texts already published, of graphical visuals of the news cycle of a certain topic in an established time-frame.

There is a large variety of applications, for commercial use and free, that can be used for Text Mining: for example, there is IBM SPSS Modeler, Leximancer, SAS Enterprise Miner, Polyanalyst (from Megaputer), and the development language and environment R. Through sophisticated computer statistical algorithms, these tools facilitate the treatment of large amounts of non-structured text data to generate scenarios, forecasts, and insights.

According to Ranfagni, Guercini e Camiciottoli (2014), market researchers have already demonstrated some interests in the use of Text Mining in consumer behavior analysis and brand awareness. In the same article, they propose a new methodological concept in which qualitative technique of digital ethnography is reinforced by the integration of quantitative techniques of Text Mining, to analyze the relationships between brands and consumers in online communities. To corroborate to the relevance of Text Mining in recent researches, the authors mention analysis that used these techniques:

- Tendency analysis: Rickman *et* Cosenza (2007 *apud* Ranfagni; Guercini; Camiciottoli, 2014) discussed the use of tools of Text Mining to analyze tendencies through tracking the "buzz" in fashion weblogs.

- Analysis of an online complaint forum: Chen (2009 *apud* Ranfagni; Guercini; Camiciottoli, 2014) used Text Mining to analyze an online complaint forum in order to determine how consumers feel, such as the main questions and reasons for dissatisfaction.

- Feelings analysis: using consumer evaluation websites as data sources, Lee *et* Bradlow (2011 *apud* RANFAGNI; GUERCINI; CAMICIOTTOLI, 2014) applied Text Mining techniques to survey positive and negative evaluations of the characteristics of digital cameras under market analysis.

- Consumer opinion analysis of electronic devices: Archak *et al*. (2007 *apud* RANFAGNI; GUERCINI; CAMICIOTTOLI, 2014) extracted business intelligence through systematic identification of words that expressed the level of satisfaction and the preferences related to characteristics of products.

- Consumer opinion analysis under automobile market: Feldman *et al*. (2008 *apud* RANFAGNI; GUERCINI; CAMICIOTTOLI, 2014), through the extraction of co-occurrence of mentions of brands and adjectives consumers used to describe opinions, the researchers studied preferences of consumers to acquire better understanding of the automobile market.

2.2 Opinion Mining

Opinion Mining is a computer study of opinions and evaluations of people related to entities, such as products, services, enterprises, personalities, and events. It is a research field in Natural Language Processing and Web Mining, which in the past few years has called attention from academic and market researchers.

Opinion Mining uses the Feeling Analysis technique, which aims to extract subjective information to determine the "poles" of specific topics, from a set of text documents, most of them online opinions expressed spontaneously by blog users, social networks, and consumer complaint websites.

The Feelings Analysis of a text segment involves the identification of the subject and of the object, as well as the evaluation of the subject regarding the object. For example, in the case of "consumer service was excellent", the object is "consumer service", and the evaluation was "excellent".

The Feelings Analysis classifies feelings in one of the three categories: positive, negative, or neutral. This research considered only positive and negative feelings. According to Nascimento *et al*. (2012), there is not a consent in literature regarding the typical characteristics of texts classified as "neutral". Many articles ignore the category "neutral", because it is difficult to find a comment that does not express any opinion, and even they are there, they are irrelevant for the analysis.

For Kim *et* Jeong (2015), Topic Extraction and Buzz Analysis usually are related to the analysis of market tendencies, while Feelings Analysis is commonly used for the analysis of reputation of products, services, and companies, and the reactions of consumers to marketing actions.

Mostafa (2013), after a bibliographical review, concludes that the majority of the applications of Feelings Analysis have been used in basically four areas: product evaluation, movie reviews, information extraction for political orientation, and stock market forecasting.

2.3 Keywords Extraction

Keywords Extraction, or Topic Extraction, is a technique that permits the understanding and the ordering of terminologies or sentences more presently common in a text document. The results can be used to promote insights, or in other words, previous understanding regarding the essence of large volumes of non-structured information.

During the search for keywords, the search for negative sentences or expressions can be enabled, such as "dislike". This feature is extremely valued by analysts who intend to distinguish feelings and opinions in text content, once while using as support the occurrence of words, there could come some distortions in meanings and misinterpretations, thus leading to situations such as a "dislike" be considered a positive opinion "like".

In Keywords Extraction, filters are applied to eliminate undesired words and expressions, as well as the ones with no relevance for the analysis, known as "stop words". Text Mining softwares in general already have a list of stop words, but most of them also enable the user to add more terms.

In this research, to observe the results of the Words Extraction, the researchers built a Word Cloud, also known as Tag Cloud. In the Word Cloud, the size of a terminology is related to the frequency in the texts, and as a consequence, its relevance in the content of the text documents, which means that as larger the word was present in the Cloud, it is possible to consider larger was its importance in the textual body.

3. FIELD RESEACH USED TO COMPARE

The goal of the field research, commissioned by the retailer network to the IBOPE[2] institute, consisted in getting information related to the opinion of the clients about the network retail stores, as well as to observe the strong and weak points of the company.

The investigations to gather data of the field research were performed through face-to-face interviews in three stores of the network, located in the city of Rio de Janeiro. The target audience asked to participate in the interviews were people shopping at the selected stores. The research took place from May 15th to May 21st 2014. The sample was selected randomly, composed by 216 interviews. On the sample of 216 cases, the margin of error is 7%, considering a trust level of 95%.

To compare with the results of the analysis done from Text Mining techniques, the following questions were considered more relevant during the field research:

- "When will you choose an enterprise of this commercial segment to go shopping, what do you take

2   IBOPE is a private research company that performs opinion surveys. The IBOPE's intelligence activities are ruled by ethical standards set by the Brazilian Association of Research Companies (ABEP, in Portuguese) and by the European Society of Market Research (ESOMAR). http://www.ibope.com.br/pt-br/ibope/quemsomos/Paginas/default.aspx

into consideration in the first place? And in second place? And in third place?";

- "Based on your experience with this store as a whole, using a scale from 0 to 10, in which 0 means you surely <u>would not recommend</u>, and 10 you surely <u>would</u> recommend this store to friends and family, what is the chance you would recommend this store to friends and family? You can pick any value between 0 and 10".

In Chart 3 (under Section 4) shows the five most mentioned attributes in the interviews used to answer the question "When will you choose an enterprise of this commercial segment to go shopping, what do you take into consideration in the first place? And in second place? And in third place?" "Price" was considered the most important attribute chosen by the consumers of the stores in this segment.

To analyze the question "Would you recommend us to a friend?", IBOPE employed the NPS (Net Promoter Score) methodology, used as a standard to evaluate the level of satisfaction of a client to a company.

According to NPS, the interviewer places the question of recommendation, and the interviewee must provide a grade between 0 and 10, being 0 considered as he would surely not recommend, and 10 means that he would surely recommend.

According to the answer, each interviewee is classified in one of the three segments of the Chart 1.

Chart 1. Classification by NPS methodology

Promoters	=	grade 9-10 →	Tendency to speak well
Passive	=	grade 7-8 →	Neutral (tendency not to speak)
Detractors	=	grade 0-6 →	Tendency to badmouth

The NPS index is defined by the difference between Promoters and Detractors

Source: Produced from IBOPE Intelligence (2014)

Image 1 presents the result from the compilation of researches performed in the three stores for the question "Would you recommend us to a friend?".

Image 1. NPS classification regarding the evaluation results from clients

Source: Produced from IBOPE Intelligence (2014)

Legend: Não recomendaria = Would not recommend; Recomendaria = Would recommend; Detratores = Detractors; Promotores = Promoters; Passivos = Passive

4. DATA ANALYSIS: APPLICATION OF TEXT MINING TECHNIQUES AND COMPARISON WITH FIELD RESEARCH

The text used for Opinion Mining were posts placed only by clients, extracted from the fan page of the retailer network company – comments and evaluations about: services, stores, products, and employees. All clients' posts placed from the beginning of March until the end of August 2014 were analyzed, in order to confront the field research done in May 2014. The interval was established so it would match the number of interviews performed in the field.

To analyze the textual content, two techniques of Text Mining were used: Keywords Extraction and Feelings Analysis. Text Mining was performed using the software *PolyAnalyst*, basically through the knots Keyword Extraction and Sentiment Analysis.

In the Keywords Extraction (Image 2), only the subjects were extracted, as means to survey the most relevant topics. The adjectives were removed in this moment because during Feelings Analysis already selects the most frequent adjectives in order to observe the most relevant evaluations of clients.
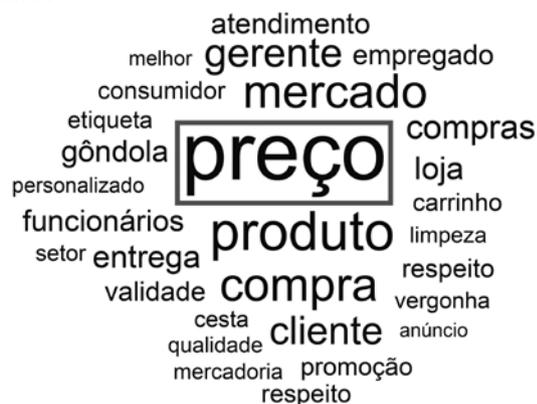

Image 2. Word Cloud generated from network's fan page

Source: The authors themselves

Legend: From the top left - Consumers' service | Best | manager | employee | consumer | market | Tag | PRICE | shopping | Aisle | store | Customized | shopping cart | Employee | product |cleanliness | Area | Delivery | respect | Shelf life | purchase | shame |Basket | client | ad | Quality |Merchandise | promotion |Respect

Comparing the results of Text Mining from the content found on the *Facebook* company's page, the most significant complaints were very similar to those results extracted from the answers of the questionnaires filled in the three physical stores of the retailer network.

Tracing a parallel between Chart 2 of Keywords Extraction from the fan page, and Chart 3, with the attributes clients interviewed in the stores consider the most

important, seen more frequently between the three first choices, there is some significant correlation between them.

Chart 2. Result of Keyword Extraction

Keyword	Relevance
Price	100%
Product	63.47%
Market	53.28%
Purchase	43.98%
Manager	33.89%
Client	30.38%
Delivery	24.58%
Employees	23.24%

Source: The authors themselves

Chart 3. Results from research in stores

	Mentioned - 1st + 2nd + 3rd places
Price	85
Variety of brands and products	64
Location	54
Environment	50
Costumers' service	46

Source: Produced from IBOPE Intelligence (2014)

In the research in stores (Chart 3), most of the clients declared "price" as their main argument for choosing the company for shopping. The result from Keyword Extraction (Chart 2) also showed the main topic mentioned by clients on company's *Facebook* page was the "price" of goods. In this same analysis, in company's fan page, "product" came as the second most mentioned topic, as in the field research, however, the later was more specific, mentioning the variety of products and brands. The third and fourth most mentioned elements in the research in the stores (Chart 3) were "location" and "environment", and some concerns related to the "market", the third most mentioned topic in Chart 2, about the discussions raised by clients at the company's social network. "Consumers' service", the fifth topic of Chart 3, in fact englobes from the fourth to the eighth topics of Chart 2 – "purchase", "manager", "client", "delivery", and "employees" – all terminologies of activities and actors involved in consumers' care.

It is important to highlight that, on charts and images generated from the analysis, either Keyword Extraction and Feelings Analysis, it is possible to select each word and contextualize it in the sentence of origin. Hence, it is possible to observe its relevance in the context it was used, analyzing the concordance to which the software is implying to the researcher.

Feelings Analysis is performed from the content found in the company's fan page (Image 3), through extraction and classification of positive and negative comments, resulting in an approval rate of 53% and disapproval rate of 47%. The research in the stores, based on the answers from clients to the question "Would you recommend us to a friend?" (Image 4), showed a disapproval rate of 45%. Comparing the two methods of analysis, the 2% difference between the results can be considered small. Therefore, again, there is a similarity in the results found.

Image 3. Result from Feelings Analysis from company's fan page
Source: The authors themselves
Legend: 47% negative; 53% positive

Image 4. Result from the field research related to the question:
"Would you recommend us to a friend?"
(extracted from Image 1 – NPS Classification)
Source: Produced from IBOPE Intelligence (2014)
Legend: Não recomendaria = Would not recommend; Recomendaria = Would recommend

Under Feelings Analysis, it was also possible to select evaluations that were more frequent in the comments at the fan page (Image 5). What was more evident was the fact that clients considered that the products of the retailer network were "very expensive".
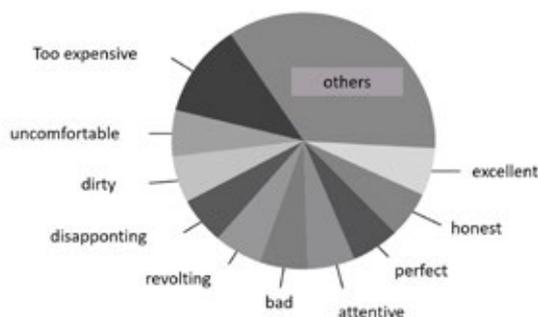

Image 5. Most frequent evaluations in Feelings Analysis
Source: The authors themselves

5. CONCLUSION

Blogs and social networks, such as *Twitter* and *Facebook*, today have a considerably important role in building the opinion of people. Then, they are also an important source for Mining and the analysis of information and opinion about many topics, such as brands, products, services, and personalities. Hence, more and more the content generated spontaneously by the users of these media convert into objects of study of areas related to the management of consumers' care, from the control of public opinion to marketing researches. Both in the academic field and in the market, the interest for market researches grows, using methods and techniques seen in Text Mining, which will support the process of transformation of a large volume of data from non-structured texts, into relevant information for decision making and strategic planning.

Based on this scenario, this present exploratory research aimed to demonstrate the advantages of the use of Text Mining techniques to learn consumers' opinions and in studies of brand perception and reputation. The effectiveness of Text Mining techniques was proven through a comparative analysis of results acquired from a field research for a retail trade network, and results found from an analysis using Feelings Analysis and Keyword Extraction of contents posted on the *Facebook* page of the company.

Feelings Analysis and Keyword Extraction achieved their main goal: the understanding of clients' opinions regarding their stores, consumers' care, and the products on sale at the network. The results found through the classification and organization, as well as the results found in the three physical stores, plus the survey of positive attributes, demonstrated the main concerns of the company seen by its clients, prioritizing questions to be discussed and lead to changes in strategy and market positioning of the retailer network.

This present work demonstrated a conformity in comparing the results found by the analysis using Text Mining techniques in content extracted from the retailer trade network fan page, with results acquired from the field research done in the same network, the exploratory research performed at the social network raised information and consumers' opinion very similar to those from the questionnaires of field research.

The results of Mining and comparison with the conclusion of personal interviews ratify the idea that Text Mining techniques are presented as effective and advantageous solutions to learn opinions and to study the perception and reputation of brands. As a consequence, they corro-

borate to the characteristics of practicality and accessibility when taking into consideration the time and the cost compared to traditional market research techniques.

Besides presenting advantages and proving the effectiveness of Text Mining techniques, this study does not intend in any way to discharge the importance of traditional market research techniques. Surely it is needed for a more objective and deeper understanding of a topic, or of one that the researcher is intending to find, using focal groups and face-to-face interviews, as traditional techniques are more appropriated for these ends, once they enable the interviewer to pursue a topic through his intervention.

As a conclusion, it is recommended to use Text Mining techniques to collect and analyze information from market research, and to analyze the opinion in brand perception studies, mainly due to the benefits regarding time and costs. However, we here propose the use of such techniques for exploratory researches in which the intention is to build a preliminary study regarding a research object, or to get familiar with the topic being investigated.

For future studies intended to apply the Text Mining techniques to extract and analyze opinions, the authors suggest a wider tracking of information from the web from comments about the company, product, or service, such as consumers' complaint websites. Besides that, to effectively increase competitive advantage, and to evaluate the business environment, companies need to monitor and analyze not only the text information present in their websites, but also in others and in the social networks of their competitors, as well as in blogs and social networks that adopt topics related to their field in the market.

REFERENCES

Camiciottoli, B.C., Ranfagni, S. & Guercini, S. (2014), "Exploring brand associations: an innovative methodological approach", *European Journal of Marketing*, 48(5), pp.1092–1112, disponível em: http://www.emeraldinsight.com/10.1108/EJM-12-2011-0770 (Acesso em 12 de junho de 2015).

Chen, L., Qi, L. & Wang, F. (2012), "Comparison of feature-level learning methods for mining online consumer reviews", *Expert Systems with Applications*, 39(10), pp.9588–9601, disponível em: http://dx.doi.org/10.1016/j.eswa.2012.02.158 (Acesso em 19 de Junho de 2015).

Chung, W. & Tseng, T.B. (2012), "Discovering business intelligence from online product reviews: A rule-induction framework", *Expert Systems with Applications*, 39(15), pp.11870–11879, disponível em: http://dx.doi.org/10.1016/j.eswa.2012.02.059. (Acesso em 8 de Junho de 2015).

Gomez, L., Mateus, C., Cardoso, H., & Rosa, C. (2013). "Territorial branding: DNA Alvito". Disponível em: http://www.ideas-revolution.pt/cms/wp-content/uploads/2013/01/IDEASpaper_ALVITO_CHILE.pdf. (Acesso em 20 de junho de 2015).

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*, disponível em: http://doi.org/10.1016/j.ijinfomgt.2013.01.001 (Acesso em 29 de Maio de 2015).

Hu, N. *et al*. (2012), "Manipulation of online reviews: An analysis of ratings, readability, and sentiments", *Decision Support Systems*, 52(3), pp.674–684, disponível em: http://dx.doi.org/10.1016/j.dss.2011.11.002 (Acesso em 24 de Julho de 2015).

IBM SPSS Modeler, disponível em: http://www-03.ibm.com/software/products/pt/spss-modeler (Acesso em 20 de agosto de 2015).

Jain, a. K., Murty, M.N. & Flynn, P.J. (1999), "Data clustering: a review", *ACM Computing Surveys*, 31(3), pp.264–323, disponível em: http://doi.org/10.1145/331499.331504 (Acesso em 7 de agosto de 2015).

Jansen, B.J. *et al*. (2009), "Twitter power: Tweets as electronic word of mouth", *Journal of the American Society for Information Science and Technology*, 60(11), pp.2169–2188, disponível em: http://doi.org/10.1002/asi.21149 (Acesso em 24 de Julho de 2015).

Khadjeh Nassirtoussi, A. *et al*. (2014), "Text mining for market prediction: A systematic review", *Expert Systems with Applications*, 41(16), pp.7653–7670, disponível em: http://doi.org/10.1016/j.eswa.2014.06.009 (Acesso em 15 de Maio de 2015).

Kim, Y. & Jeong, S.R. (2015), "Opinion-Mining Methodology for Social Media Analytics", *KSII Transactions on Internet and Information Systems*, 9(1), pp.391–406

Leximancer, disponível em: http://info.leximancer.com/ (Acesso em 20 de agosto de 2015).

Li, Y. & Li, T. (2013), "Deriving market intelligence from microblogs", *Decision Support Systems*, disponível em: http://dx.doi.org/10.1016/j.dss.2013.01.023 (Acesso em 17 de junho de 2015).

Liu, H. *et al*. (2013), "Electronic Commerce Research and Applications Combining user preferences and user opinions for accurate recommendation", *Electronic Commerce Research and Applications*, 12(1), pp.14–23, disponível em: http://dx.doi.org/10.1016/j.elerap.2012.05.002 (Acesso em 12 de Junho de 2015).

Loh, S., Wives, L.K. & de Oliveira, J.P.M. (2000), "Concept-based knowledge discovery in texts extracted from the Web", *ACM SIGKDD Explorations Newsletter*, 2(1), pp.29–39, disponível em: http://dl.acm.org/citation.cfm?id=360414 (Acesso em 20 de Fevereiro de 2015).

Megaputer - PolyAnalyst, disponível em: http://www.megaputer.com/site/polyanalyst.php (Acesso em 20 de agosto de 2015).

Milolidakis, G. & Kimble, C. (2013), "Digital traces for business intelligence. A case study of mobile telecoms service", Journal of Enterprise Information Management, 27(1), pp. 66-98, disponível em: http://www.emeraldinsight.com/doi/full/10.1108/JEIM-09-2012-0061 (Acesso em 27 de Fevereiro de 2015).

Mostafa, M.M. (2013), "More than words: Social networks' text mining for consumer brand sentiments", *Expert Systems with Applications*, 40(10), pp.4241–4251, disponível em: http://dx.doi.org/10.1016/j.eswa.2013.01.019 (Acesso em 20 de Fevereiro de 2015).

Nascimento, P. *et al*. (2009), "Análise de sentimento de tweets com foco em notícias", disponível em: http://www.lbd.dcc.ufmg.br/bdbcomp/servlet/Trabalho?id=11993 (Acesso em 6 de março de 2015).

Poria, S. *et al*. (2014), "EmoSenticSpace: A novel framework for affective common-sense reasoning", *Knowledge-Based Systems*, disponível em: http://dx.doi.org/10.1016/j.knosys.2014.06.011. (Acesso em 10 de maio de 2015).

R Project for Statistical Computing, disponível em: https://www.r-project.org/ (Acesso em 20 de agosto de 2015).

Ranfagni, S., Guercini, S. & Camiciottoli, B.C. (2014), "An interdisciplinary method for brand association research", *Management Decision*, 52(4), pp.724–736, disponível em: http://www.emeraldinsight.com/10.1108/MD-04-2012-0284. (Acesso em 20 de Julho de 2015).

Reyes, A. & Rosso, P. (2012), "Making objective decisions from subjective data: Detecting irony in customer reviews", *Decision Support Systems*, 53(4), pp.754–760, disponível em: http://dx.doi.org/10.1016/j.dss.2012.05.027. (Acesso em 20 de abril de 2015).

SAS Enterprise Miner, disponível em: http://www.sas.com/en_us/software/analytics/enterprise-miner.html. (Acesso em 20 agosto 2015).

Seol, H., Lee, S. & Kim, C. (2011), "Identifying new business areas using patent information: A DEA and text mining approach", *Expert Systems with Applications*, 38(4), pp.2933–2941, disponível em: http://doi.org/10.1016/j.eswa.2010.06.083 (Acesso em 5 de Fevereiro de 2015).

Sobkowicz, P., Kaschesky, M. & Bouchard, G. (2012), "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web", *Government Information Quarterly*, 29(4), pp.470–479, disponível em: http://dx.doi.org/10.1016/j.giq.2012.06.005 (Acesso em 14 de Março de 2015).

Thorleuchter, D. & Poel, D. Van Den (2012), "Predicting e-commerce company success by mining the text of its publicly-accessible website", Expert Systems with Applications, 39, pp.13026–13034, disponível em: http://dx.doi.org/10.1016/j.eswa.2012.05.096 (Acesso em 15 de Março de 2015).

Xu, K., Guo, X., *et al*. (2012), "Electronic Commerce Research and Applications Discovering target groups in social networking sites: An effective method for maximizing joint influential power", *Electronic Commerce Research and Applications*, 11(4), pp.318–334, disponível em: http://dx.doi.org/10.1016/j.elerap.2012.01.002 (Acesso em 6 de Junho de 2015).

Xu, K., Li, J. & Song, Y. (2012), "Identifying valuable customers on social networking sites for profit maximization", *Expert Systems with Applications*, 39(17), pp.13009–13018, disponível em: http://dx.doi.org/10.1016/j.eswa.2012.05.098 (Acesso em 12 de Julho de 2015).

Yu, Y., Duan, W. & Cao, Q. (2013), "The impact of social and conventional media on firm equity value: A sentiment analysis approach", *Decision Support Systems*, disponível em: http://dx.doi.org/10.1016/j.dss.2012.12.028 (Acesso em 3 de Fevereiro de 2015).

Zerr, S. *et al*. (2013), "Sentiment Analysis and Opinion Mining in Collections of Qualitative Data", http://www.arcomem.eu/wp-content/uploads/2013/08/ipres_gute_arbeit.pdf (Acesso em 11 de Abril de 2015).

Zhang, L. & Liu, B. (2014), "Aspect and entity extraction for opinion mining", *Data Mining and Knowledge Discovery for Big Data*, pp.1–40, disponível em: http://doi.org/10.1007/978-3-642-40837-3 (Acesso em 14 de Maio de 2015).

Zubrinic, K., Kalpic, D. & Milicevic, M. (2012), "The automatic creation of concept maps from documents written using morphologically rich languages", *Expert Systems with Applications*, 39(16), pp.12709–12718, disponível em: http://dx.doi.org/10.1016/j.eswa.2012.04.065. (Acesso em 11 de Abril de 2015).