Sentiment analysis in social networks: a study on vehicles

Renata Maria Abrantes Baracho,
Gabriel Caires Silva, Luiz Gustavo Fonseca Ferreira

[1]Progama de Pós-Graduação em Ciência da Informação (PPGCI)
Universidade Federal de Minas Gerais (UFMG)
P.O. 486 31270-901 – Belo Horizonte – MG – Brazil

Abstract. *This paper presents partial results of a research project that aims to create a process of sentiment analysis based on ontologies in the automobile domain and then to develop a prototype. The process aims at making a social media analysis, identifying feelings and opinions about brands and vehicle parts. The method that guided the development process involves the construction of ontologies and a dictionary of terms that reflect the structure of the vocabulary domain. The proposed process is capable of generating information that answers questions such as:* "In the opinion of the customer, which car is better: Corsa or Palio? Which one is more beautiful? Which engine is stronger?" *To answer these questions by comparison, one can show a general view reflected on different social networks, indicating, for example, that for a given vehicle, a certain percentage of responses are considered positive, while for others, the percentage is considered negative.*
The results can be used for various purposes such as guiding decisions to improve the products or directing specific marketing strategies. The process can be generalized and applied to other areas in which organizations are interested in monitoring views expressed about their products and services.[1]

1. Introduction

The increase of personal information available on the Web, especially in recent years, is noteworthy at least. With the advent of what is called *Web 2.0*, countless opinions and feelings about every subject, are wildly available throughout the Web. In this new era, besides the content offered by companies and organizations, individuals have come to share reviews and opinions via personal blogs, networking sites, and microblogs, just to name a few.

This paper presents the initial results of a research project whose main objective is to create a model of knowledge representation in the context of social networks on the Web. In this project we developed a prototype software for sentiment analysis of an automobile brand on the Web. This is archived by the use of morphologic analysis, and language features detection aided by ontologies. Specific objectives include the design of methodologies for opinion mining, composition and classification, creation of a dictionary of terms that contains sentiment orientation by translating this type of dictionary from another language, design and use of ontologies to be used in the process of sentiment detection and data summarization and finally the working prototype itself.

[1]This work is partially supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Governo do Estado de Minas Gerais, Brazil, Rua Raul Pompéia, nº101 - São Pedro, Belo Horizonte, MG, 30.330-080, Brazil.

Our prototype is applied to a specific company in the automobile market (FIAT) and presents innovative nature of monitoring business intelligence and user opinion. It targets information available on the Web from several sources (such as automotive centered portals, blogs and discussion groups) in Portuguese language, although the presented methodology could be easily applied to any other language, source or targeted domain. The objective for the prototype is that it should be able to answer questions and give important insights about the sentiment on car brands on the Web. For example: *"what are people saying about FIAT Punto in social media?"* These results can be used to improve products or direct marketing strategies, as well as be applied by any other organization interested in monitoring sentiment about a product and/or service.

The proposed methodology and our resulting prototype collects, structures and analyzes Web information by using a combination of text processing technologies with several other linguistic techniques such as morphological, syntactic and semantic analysis guided by the target domain terms supplied by our ontologies and the list of terms that evokes sentiment (with a given polarity and strength value) supplied by our sentiment dictionary. The information in our ontology is structured as trees of objects that relate to each other as a *is part of* or *is one* relationships. Each object may have one or more terms that can be used to identify references to that particular object on the sentences extracted.

Each sentiment detected in the process described is then stored in the ontology tree structure, as a link of one or more objects in the ontology tree to a sentiment value, that can be either neutral, positive or negative. Using this hybrid unsupervised approach, by combining language processing, lexicon techniques and ontology techniques for the sentiment data structure, we expect to generate classifiers for sentiments and opinions and business intelligence insights that improve the results obtained so far in sentiment analysis and opinion mining without relying on supervised algorithms such as machine-learning approaches that requires a costly training phase that may be impeditive for groups with limited resources.

This project is funded by the state financing agency. The group is composed of two PhD and two graduate researchers. The project's main area of research is Information and Knowledge Management, but as our main objective is to retrieve information from texts in natural language processing them lexically and morphologically to extract semantic proprieties (sentiments) with the use of ontologies we also relate the areas of Information Extraction and Retrieval, Knowledge Representation, Analysis of Information Systems, Semantic Technology and Philosophy of Information.

2. Theoretical Framework

With the rise of Web 2.0 and specialized portals, blogs, and social networks, an enormous amount of new personal opinion is made accessible on a daily basis. Reviews, ratings, recommendations and other forms of expression are available on-line. Information previously obtained through a costly and time consuming process of satisfaction and opinion research can now be obtained on a large scale on the Web. The new challenge is how to process and interpret this massive amount of information, and this challenge is the object of research in the discipline called "sentiment analysis and opinion mining". In this research we consider the following definition for the term "sentiment analysis and opinion mining:" the identification, extraction and study of opinions, feelings and emotions expressed in texts from the web. In the following section, the theoretical framework of the research project can be found. The research papers reviewed below inspired the developments of the method and process of sentiment analysis detailed in Section 3.

2.1. Ontologies

[Cicortas et al. 2009] There is a growing demand for information systems-oriented interpretation of human language. These systems are designed to be capable of understanding the intentions and opinions of the author with minimal human intervention. In the article entitled Considerations on Ontologies Construction, the author identifies the challenge the interpretation of heterogeneous information by automated tools and analyzes possibilities of using ontology to resolve these issues. The combination of ontological and natural language rules are seen as a solution to improve performance of sentiment analysis.

Also in this context, the importance of ontologies in identifying the meaning of information, through detailed description of complex systems is highlighted, [Rösner and Kunze 2003]. Also discussions about the best practices for building ontologies. The authors present their experiences related to construction of new ontologies, detailing various methods for the use of language constraints, design principles and ideas for frameworks. They emphasize the importance of having a quality system to detect synonyms in a process of creating ontologies, since its absence in many cases, can compromise the quality of the results.

[Polpinij and Ghose 2008] In an article *An Ontology-Based Sentiment Classification Methodology for Online Consumer Reviews*, classification is presented as a proposed ontological approach based on lexical variation. The authors propose the use of three sources for the construction of an ontology: a dictionary, a list of text and a set of verbs. From these sources the ontology is built based on three types of information: morphological analysis (indicating a pattern in the composition of the word), a parse (containing information about their classification, e.g. verbs and suffixes such as e.g. and e.s.), and finally a semantic analysis based on logical constraints of synonymy, antonymy and subsumption (relationship “is one”).

The ontological structure derived is then used to create a model BOW (“bag of words”) and fed into a classifier. According to the authors, this technique achieved satisfactory results, reaching 96% accuracy. [Kunze and Rösner 2005] present a methodology for ontology extension using concepts derived from a specific domain. The method uses a first and a body ontology partially processed in the domain. The approach is based on syntactic and grammatical structures and basically explores features of the language contained in the input corpus.

2.2. Sentiment Analysis

[Liu 2010] presents an introduction to key problems and solutions within the existing area of sentiment analysis research highlighting its importance both to individuals and to companies in market research and interest / customer satisfaction. The text provides important definitions such as the concepts of object and features (properties or parts of an object) and opinion (feeling positive, negative or neutral in relation to an object or a feature).

[Wang et al. 2011] propose a method of selection of features for the classification of feelings. Based on linear discriminant analysis (Fisher's discriminant ratio), the method utilizes the concept of information gain (Information Gain) and is validated through the comparison with other methods based on this concept. In the article, the authors present the results of two experiments in which selection methods tested different features. The experimental results indicated that the linear discriminant method has better performance than the others analyzed.

The approaches used by [Ramanathan and Ramnath 2010] explore the use of context in sentiment analysis using three techniques. The first is an approach that makes use of domain ontology mapping sentences on objects in the ontology. For each object, a weight is defined positive and negative and the positive and negative score of a sentence is defined as the sum of these weights. The weights are defined using machine learning techniques and regression. The second approach makes use of a technique for capturing sequences of characters that appear frequently. For each pair of sentences, extracted a set of words that appear in both, and each of these sets, you assign a score positively or negatively according to how often they appear in sentences. Finally, two approaches are used to combine three techniques for the classification of polarity of a sentence and the results presented.

[Wei and Gulla 2010] present analysis technique based on a tree of feelings of ontological features. The tree SOT (Sentiment Ontology Tree) is constructed to represent the features of an object hierarchy. Each node the tree contains as children. Besides these features there are two leaf nodes representing the negative and positive feelings of the feature represented by the node. The classification approach used is based on hierarchical classification algorithm. The algorithm takes as input a SOT and texts already sorted and aims to validate the hierarchical construction of sentimental texts. The results demonstrated that knowledge of hierarchical relations improve performance and accuracy of sentiment analysis. In addition, you can use a generic model with a SOT composite, SOT of individual objects, and a root node. This adjustment allows the algorithm to be used with general texts (i.e. not containing a predefined object).

[Neviarouskaya et al. 2011] article in The Lexicon for Sentiment Analysis describes a method for automating the generation and marking values for level of feeling subjective text fragments called SentiFul. The idea is to enable any basis to expand through techniques such as direct synonymy, antonym, relations of exploitation, hyponymy derivation, and composition, among others. The proposal is made pursuant to textual recognition, using four types of affixes (used in the derivation of new words), depending on their role with regard to feelings such as propagation, reversal, intensification, and weakening. The derivation is done to find new words using such composition. This process generates a large number of terms useful especially in the case of nouns and adjectives. The algorithm is designed for the automatic extraction of words related to sentiment using terms from WordNet (but using words from SentiFul).

3. Tools and methods

This section presents the tools and methods that were used in the analysis process of feeling as well as a detailed description of the process developed and proposed in this research.

3.1. Tools for the process of sentiment analysis

Below the software that makes the PALAVRAS software which performs the semantic analysis of text and was used in the development of the process and the creation of the feelings dictionary is described.

3.2. The Palavras software

The process developed in the research consisted primarily of semantically analyzing fragments of texts (articles and reviews) of social networks in order to extract information from feelings. To do a semantic analysis of text using WORDS software (developed by Eckhard Bick and based on corpus "Syntactic Forest" of Linguateca) was performed. This is an automatic parser for Portuguese that performs parsing, syntax analysis of the Portuguese language and is able to provide morphological information of a sentence.

The process of sentiment analysis prepared by the research begins with the use of WORDS as parser and lexicon. This software is used as the basis of the algorithm, "normalizing" the input and parser. The process of sentiment analysis begins with the extraction of text elements related to the view, then uses the classification of opinion as to his character considered within the scope of positive, negative or neutral. The sequence is performed to compare their opinions and judgments, and commonly uses the term "object" to refer to the target's opinion, which may contain several features or subparts. These may also be subject to reviews.

3.3. The sentiment dictionary used

The construction of the dictionary of feelings was based on the classification of feelings dictionary Sentistrenght [Thelwall et al. 2012], a dictionary with tens of thousands of terms denoting feelings. The purpose of the dictionary is to quantify feelings. A sample entry in this dictionary would be: "bad: -2", which means that the English word carries a negative feeling with numeric value -2. According to the authors, the dictionary was constructed from research in psychology, philosophy and linguistics. The biggest challenge for the use of the dictionary SentiStrenght during the project was to undertake the translation process while maintaining the real meaning of the words in the English language. The process of translation dictionary Sentistrenght that contains a list of approximately over 22,000 words, through a process of semi-automatic translation, was divided into three steps described below. The first step comprises the initial translation.

We used three tools: the Bing Translator from Microsoft, Google and Yahoo BabelFish Translator. From the translation made by each tool, an index of agreement was created. With this it was possible to filter terms with higher disagreement among dictionaries and therefore need more attention. The second step consists in validating the translation of terms. Despite the undoubted utility of translators, automated much texture characteristics of each tongue are not detected, so that often the manual intervention of a person skilled in translation is necessary. As it would be very costly to allocate an expert to translate all the terms in question, it was decided to automate the process. The process consists in the execution of a program designed to access the COMPARA, which is a parallel corpus English/Portuguese available from Linguateca (distributed resource center for the Portuguese language). The operation of COMPARA is as follows: given a term in English (or Portuguese), he shows us how it was the translation of that term in several different contexts, including works of Machado de Assis, Eca de Queiros and Aluisio Azevedo.

The program developed at this stage serves as a crawler, referring COMPARA for each of the search terms and registering cases where translations are relevant and where there is no match. Thus, we validate the suggested translation of automated translators from translations made by professionals. The third step was termination. At this stage we select the most relevant terms (i.e., terms that would result in inaccurate translations and have greater negative impact in the search results). With these terms, an inspection was made in each of the translations, looking for imperfections in the translation process done so far. In this step, just under 600 terms were analyzed.

3.4. Domain-specific ontologies

After defining the analysis of texts and the process of translation dictionaries of terms related to feelings that are considered in the research, we moved to the step of defining ontologies. The concept of ontology was used to guide the process of identifying objects (in this case car models of Fiat and non-FIAT), characteristics (properties of objects such as power, beauty, etc.) , contexts (following paragraphs of opinions on the same object). These are also vital in the final classification. For the development of this research, we used the concept of ontology used in the field of information science as a formal knowledge, a set of concepts, and their relationships in a domain. Ontologies can be used to model human and abstract concepts. The formalization allows systems to take advantage of human models. Ontologies were created based on models of FIAT. The first ontology of "FIAT models" used in this research adds the necessary information about the Fiat car models. This ontology allows us to not only detect objects of interest in the text (e.g. "Fiat Palio") but also to determine the class that the given object belongs to (a car of the "Hatch" subtype or "Hatch Compact" as shown in Figure 1.

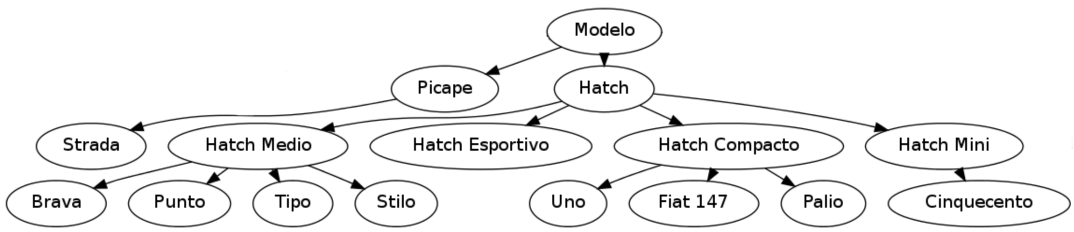

Figure 1: Example of part of the ontology: FIAT Models

The second ontology "Car Features" organizes the relevant features such as an automobile, or adds components and features that are targets of feelings (Figure 2).

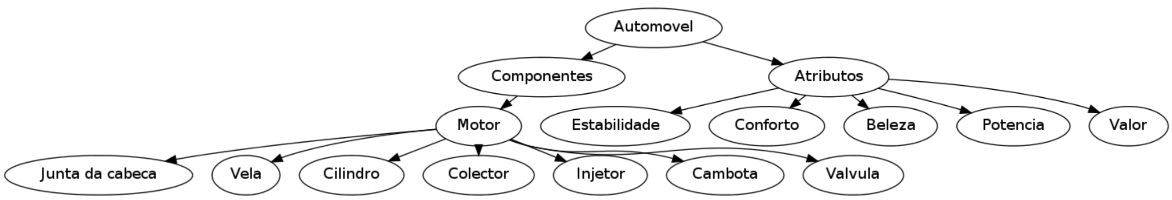

Figure 2: Example of part of the ontology: Features of a car

The third ontology "Non-FIAT Models" is made up of other objects (cars) that have the same features (features and components) of the objects of interest. The features associated with these non-FIAT objects must be detected and correctly excluded from the analysis because it would cause distortion of the result.

3.5. Complete overview of the analysis process

After the definition of ontologies the process of sentiment analysis proposed was implemented in a prototype. Generally, the process consists in the capture, analysis and storage of opinions. More specifically, the process is divided into eight stages that come from the collection of opinions on social networks to the aggregation of the opinions rendered, as shown in Figure 3. The first steps to represent the views containing the text is captured and standardized. Then the objects of interest (defined by the ontology) are found in the text, as well as their features (characteristics). Subsequently, the detection and the calculation (based on dictionary feelings) of sentiment related to each of the objects of interest and its features are done. Finally, the results are analyzed and stored in forms of reports.

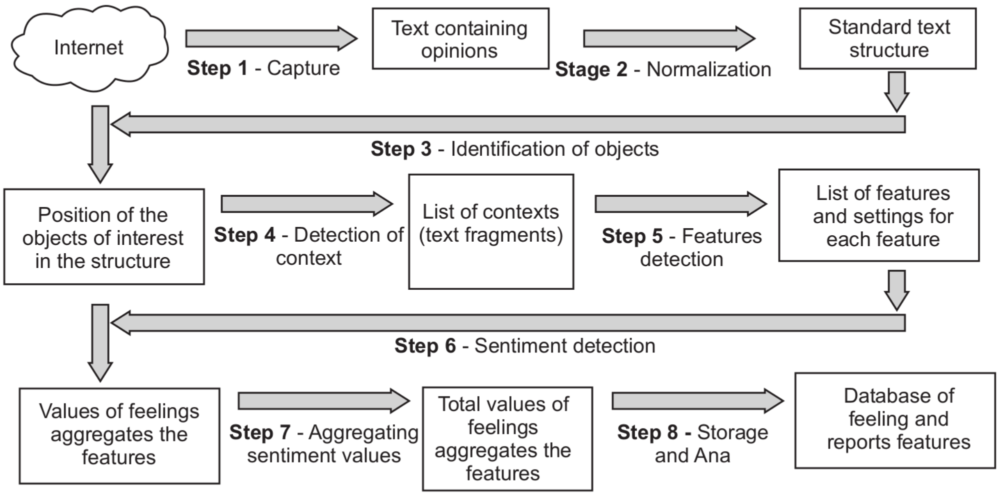

Figure 3: Overview of the review process.

3.6. Step 1 - Capture

Initially, several blogs themed "cars", which were openly discussed among readers were accessed by a crawler to build the basis of texts (corpus analysis). A funding program was developed to follow accompany the new articles and news sites selected, via RSS (RDF Site Summary). The capture process started by selecting RSS feeds (RSS news aggregator) from sites of interest, such as portals and blogs with the car theme. Then the inner content is extracted and stored as raw XML data and HTML (RSS formats in which are stored on the web). In the second step these raw data are processed by separating the HTML and XML texts of interest to the news and any comments. The result of this step is a collection of texts grouped as news/article (title and text) and the texts of any related comments. This corpus comprises the input for normalization algorithm.

3.7. Stage 2 - Normalization

After the texts were captured, it was necessary to normalize and prepare themfor algorithm review. In order to this, it is necessary to obtain the morphological and lexical structure of the texts, and bring all terms to their most basic form (infinitive). This step is performed by PALAVRAS software. Below is an example of normalization using the words in a text captured (Figure 4): "I do not know if it's the best, but I am very pleased with my Palio. It is economical and never gave me problems and I have it already since 2003." Text taken from `http://br.answers.yahoo.com/question/index?qid=20061112175518AAfs7a3`.

não [não] **ADV** @ADVL>
sei [saber] <vt> <fmc> **V** PR 1S IND VFIN
@FMV
se [se] **KS** @SUB **@#FS-<ACC**
é [ser] <vK> **V** PR 3S IND VFIN @FMV
o [o] <artd> **DET** M S @>N
melhor [bom] <KOMP> <SUP> <n> **ADJ** M
S @<SC
,
mas [mas] **KC** @CO
estou [estar] <vK> <fmc> **V** PR 1S IND
VFIN @FMV
muito [muito] <quant> **ADV** @>A
satisfeito [satisfazer] <vt> **V** PCP M S
@<SC
com [com] **PRP** @<ADVL
o [o] <artd> **DET** M S @>N
meu [meu] <poss 1S> **DET** M S @>N
pálio [pálio] **N** M S @P<
.

Figure 4: Normalization output.

3.8. Step 3 - Identification of objects.

Along with the standardized texts, in this step the ontologies of the Fiat brand cars (besides the very word FIAT) and the ontology of brands "Non-FIAT" is used to perform the detection of objects of interest. This step results in the positions of words in texts identified as objects of interest, being the objects of the models identified both FIAT cars (car models identified by Fiat) and non-FIAT (models from other companies) in the texts, we call these positions simply markers of objects that represent the position of words in the text, identified as objects of interest (car models). Recording the positioning is done by counting the number of words required to reach the word object as the beginning of the phrase or the piece of text, that is, the first word reading the text position is zero, the second is a position, and so on.

For example, the text is the sentence: "I prefer the Corsa, it is softer and more comfortable. The Palio is also good, and has more interior space. But the Gol, I think is very hard." removal of `http://br.answers.yahoo.com/question/index?qid=20070227083007AAf9lKv`. This text contains an object of interest present in the FIAT ontology (Palio object) at position 10, and two other objects of the Non-FIAT ontology (the Corsa objects at position 3 and Gol position number 21).

3.9. Step 4 - Detection of context

The next step consists of extracting the contexts of the objects detected in the previous step. A context, in the proposed process, is represented by a piece of text. Thus, the text of the sentence example of Step 3 is divided into contexts, as illustrated below (Figure 5).

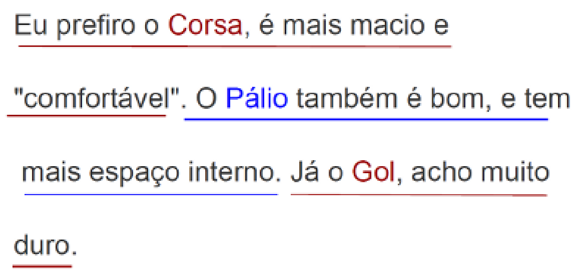

Figure 5: Context detection. "I prefer the Corsa, is softer and more comfortable. The Palio is also good, and has more interior space. I think the Goal is too cumbersome."

3.10. Step 5 - Features detection

Separate contexts, the next step in the detection takes place in the nature or subparts (features) of the car in each of the contexts generated in the previous step. In this case we, use the ontology features of cars (Figure 2) to guide the process (Figure 6).

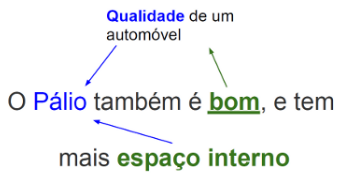

Figure 6: Features detection.

3.11. Step 6 - Sentiment detection

After detecting the features, the sentiment detection process is carried out. At this stage, SentiStrenght is used to detect and classify the sentiment level. Feelings are related features closer (Figure 7).

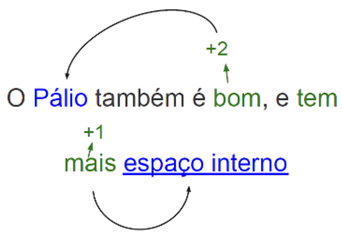

Figure 7: Sentiment detection.

3.12. Step 7 - Aggregating sentiment values

Then it is made of an aggregate structure of the detected feelings ontology feature cars (Figure 2). It can therefore be inferred, for example, if somebody speaks well about the power of a car, he/she is referring indirectly to the motor of the car. When checking the feeling at any point on the ontology tree , the feeling values aggregated to the descendants at that point are also recorded for the parent, as well as for all the other points up to it. the feelings of the current point. Therefore a positive feeling about the tires of a car model, is automatically recorded to the car itself reaching the car brand car that gets all the sentiment at the root of the ontology tree.

3.13. Step 8: Storage and Analysis

At this stage of feeling all the information is stored, keeping references to the car model (FIAT ontology), feature (cars ontology) and frameworks for future validations. This allows various types of cross-references and comparisons.

4. Results

The first result obtained in this project was the construction of the ontology of objects of interest of Fiat cars, the brand researched, the nonFiat objects, and components of cars. These ontologies represent concepts and relationships within the domain of the car market. From these data models can make inferences about the domain objects. Were used both for identifying our target objects, as a bag of words (*BOW*) initially and latter as a sentiment holder linking objects to sentiments. This allowed the system to measure sentiments at any level of the ontology structure by summarizing the sentiment assigned to descendant objects.

Our second result, in a smaller scale, is our sentiment dictionary. Our methodology was successful in translating this type of dictionary from another language without losing too much consistency and meaning while keeping the sentiment strength values. Although some human labor was needed to check the word list coherence most of the process was automatized and could be applied to other languages as well that lacks or have little availability to a proper sentiment dictionary.

The proposed methodology to extract, process and analyze sentiments and opinions on the Web, by using a unsupervised algorithm and the resulting prototype created by implementing this methodology are our third and most important result. When properly loaded with processed sentiment and relationship information, ontology trees can provide a dynamic yet simple way to relate, summarize and visualize processed data. Based on that information, many types of reports can be generated without the need of reprocessing data or sentiments. Comparative studies are a very good example of this concept. With a loaded ontology tree, important questions can be answered such as "Which car is better, based on customer feedback, Corsa or Palio?". The algorithm just needs to summarize every sentiment related to Palio and Corsa, and the decedents. But if the question is 'Which car is considered prettier?", then just sentiments related to each car and its appearance (remembering that appearance is an ontology object defined as *part of* a vehicle) are used. Many different views on the ontology that answers countless meaningful questions like those on consumer opinions, can be generated.

Figure 8 shows our results on the test made partially implementing our methodology that can serve as an example of the report output result that can be obtained from the prototype. Keep in mind that this is a preliminary result, just to exemplify what kind of report can be made using our methodology. This was run as a profofconcept to test if our methodology could process a simple dataset and provide a useful (even if limited) report. We expect much more detailed tests in the following months.

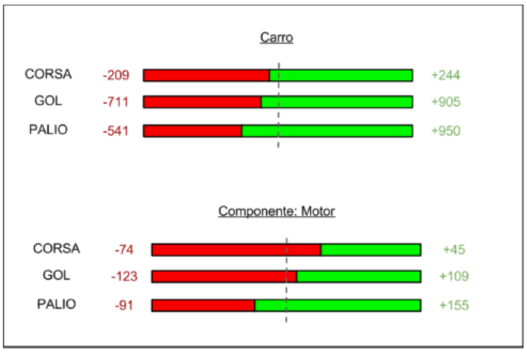

Figure 8: Partial results.

In this particular case (Figure 8 results), opinions were extracted only concerning *Corsa*, *Gol* and *Palio*, which are car brands in direct competition in Brazil. The dataset was obtained from open articles found in social media on the Web mainly constituted by news feeds from many automotive magazines sites, blogs and discussion groups. This dataset comprised 8643, articles with 83.571 related comments, for a total of 607.8527 words collected in 53.4846 sentences written by 4.112 authors (combining articles and comments data). Based on the results obtained it can easily be inferred that there are more positive than negative opinions about Palio: 63% of the opinions analyzed were considered positive while 37% were considered negative. For the brand Gol, 47% of the opinions were positive, while 53% were negative. Finally Corsa 38% positive opinions while 62% were negative for the prototype. It can be seen that, while Gol showed a more mixed public opinion result, having almost the same amount of positive and negative opinions, Palio had a more positive public opinion Corse a more negative one.

The validation of results for this test run is still ongoing and, since it had to be done by hand, only a very small set of data (only 24 articles and 63 comments) until now. This represents yet a too small validation information to provide meaningful confidence intervals. However in the set already validated we were able to perceive small number of wrong sentiment values or false positive ones (less than 6%) but, a considerably number of false negatives or sentiments not recognized (close to 15%) that we believe can be attributed to the very simplistic and small ontology trees that were used in this first test.

5. Concluding remarks

This paper presented a methodology to process, analyze and summarize sentiments and opinions from sentences extracted from the Web. This methodology was applied to the automotive field for specific analysis of the FIAT brand, resulting in a prototype that is still incomplete, but which could easily be applied to any other domain. Although the practical results were very simple, especially considering the potential for the methodology, it was sufficiently successful to serve as a proofofconcept that the methodology works and can provide interesting insights about the data.

Our practical result, on Figure 8 is an illustration of how we can visualize results from the methodology. As our prototype is in its early stages of development this result is just a proof of concept as only a small dataset was processed for just 3 objects of the ontology trees: Corsa, Palio and Gol brands of vehicles. As the development progresses we will be able to show much more detailed insights of much larger datasets. In this particular result, we showed that the Palio brand received a much more positive sentiment value then negative, while Corsa brand received the opposite results and Gol opinions were more mixed.

In a recent conference the group was asked about the possibility of changing the process described in Section 3 to make it able to collect information about the author's identity details (sex, age, country) and if the reports could be placed on a timeline, so that analysis could also identify trends. These ideas were noted as possible future developments since it could enrich the scope of the methodology described here. Also as future work we intend to generalize the proposed process so it could be even more easily applied to other areas and organizations interested in monitoring opinions expressed about products and services. A first idea to implement this generalization would be using a SOT tree technique proposed by [Wei and Gulla 2010] in lieu of our simplistic ontology tree.

References

Cicortas, A., Iordan, V., and Fortis, A. (2009). Considerations on construction ontologies. *CoRR*, abs/0905.4601.

Kunze, M. and Rösner, D. (2005). Context related derivation of word senses. *CoRR*, abs/cs/0501095.

Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing*, 2nd ed.

Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2011). Sentiful: A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing*, 2:22–36.

Polpinij, J. and Ghose, A. K. (2008). An ontology-based sentiment classification methodology for online consumer reviews. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*, WI-IAT '08, pages 518–524, Washington, DC, USA. IEEE Computer Society.

Ramanathan, J. and Ramnath, R. (2010). Context-assisted sentiment analysis. In *The 25th Annual ACM Symposium on Applied Computing*, pages 404–413, Sierre, Switzerland. ACM.

Rösner, D. and Kunze, M. (2003). Exploiting sublanguage and domain characteristics in a bootstrapping approach to lexicon and ontology creation. *CoRR*, cs.CL/0304035.

Thelwall, M., Buckley, K., and Paltoglou, G. (2012). Sentiment strength detection for the social web. *JASIST*, 63(1):163–173.

Wang, S., Li, D., Song, X., Wei, Y., and Li, H. (2011). A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. *Expert Syst. Appl.*, 38(7):8696–8702.

Wei, W. and Gulla, J. A. (2010). Sentiment learning on product reviews via sentiment ontology tree. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10, pages 404–413, Stroudsburg, PA, USA. Association for Computational Linguistics.