A Spatial and Temporal Sentiment Analysis Approach Applied to Twitter Microtexts

André Luiz Firmino Alves[1,2], Cláudio de Souza Baptista[1], Anderson Almeida Firmino[1], Maxwell Guimarães de Oliveira[1], Anselmo Cardoso de Paiva[3]

[1] Universidade Federal de Campina Grande, Brazil
baptista@dsc.ufcg.edu.br, anderson.firmino@ccc.ufcg.edu.br, maxwell@ufcg.edu.br
[2] Universidade Estadual da Paraíba, Brazil
andre@uepb.edu.br
[3] Universidade Federal do Maranhão, Brazil
paiva@deinf.ufma.br

Abstract. The widespread of social communication media in the Web has produced a large volume of opinionated textual data stored in digital format. Social media constitutes a rich source for sentiment analysis and understanding of the opinions spontaneously expressed. Many scientific proposals have arisen in the last years aiming to deal with sentiment analysis issues. However, most of them do not address both spatial and temporal dimensions that enable a more accurate analysis and a better understanding of the mood of people when using social media. In this context, we rely on Geographical Information Retrieval techniques in order to infer geographical locations mentioned in Twitter microtexts (tweets). We propose an approach based on two well-known classification algorithms for detecting the sentiment polarity on tweets considering both spatial and temporal information. Our approach differs from related work since it does not rely on Part-Of-Speech (POS)Taggers. The proposed approach is evaluated through a case study using a dataset of Portuguese Twitter microtexts harvested during a big event which took place in Brazil. The achieved results not only outperformed related work as they have shown which is possible to perform sentiment analysis with a good accuracy even without relying on POS-Taggers

Categories and Subject Descriptors: H.2.8 [**DATABASE MANAGEMENT**]: Database Applications—*Data Mining*; H.3 [**Information Storage and Retrieval**]: Miscellaneous; I.7 [**Document and Text Processing**]: Miscellaneous

Keywords: Geographical Information Retrieval; Machine Learning Techniques;Spatial; Temporal Sentiment Analysis

1. INTRODUCTION

The Web 2.0 has led to a widespread of non-structured information by users in blogs, discussion forums, online product evaluation sites, microblogs and several social networks. This fact has brought out new challenges and opportunities in the information retrieval research field [Eirinaki et al. 2012].

It is important to notice that the opinions about several themes expressed by Web users are made in a spontaneous manner and in real time [Calais Guerra et al. 2011]. In this context, sentiment analysis has emerged providing the possibility of capturing opinions of the general public, in an automated way, concerning given themes. This research field has an increasingly interest to both the scientific and the business communities. It is an open field with many research issues and challenges, due to the benefits of understanding the mood of people instantly and automatically.

With this new way of using the Web, users do not simply browse it; they actively contribute to its contents through applications, helping to build a collective intelligence [O'Reilly 2007]. Such intelligence has spread to several domains, especially those related to people's daily life, such as commerce, tourism, education and health, causing the social Web to expand exponentially [Appelquist et al. 2010]. The understanding of what people are thinking or their opinions is fundamental for the decision making process, mainly in the context that people express their comments voluntarily with the aim to cooperate to each other.

According to Liu [2012], sentiment analysis, also known in the literature as opinion mining, is the study field that analyses people's sentiments, evaluations, attitudes and emotions concerning entities such as products, services, organizations, individuals, issues, events, topics and their attributes. Sentiment analysis is a recent research field which combines advanced techniques from Text Mining, Machine Learning, Information Retrieval and Natural Language Processing (NLP) to process large amounts of non-structured content generated by users, mainly in social media [Sharma and Dey 2012]. The main purposes of sentiment analysis include the extraction of opinion and the subjective knowledge from online texts; and the formalization of this discovered knowledge in order to enable such knowledge to be analyzed for specific purposes [Liu 2012].

Unfortunately, the analysis of opinionated comments expressed in social media requires too much effort to be carried out manually, mainly due to the big volume of data. Hence, we seek for a summarization of opinions. A common way of accomplishing this summarization is by means of the classification of the opinion of an object into categories: positive, negative and neutral. This kind of classification is referenced in the literature as sentiment polarity or polarity classification [Liu 2012].

Twitter is a rich source to understand people's opinions about many aspects of their daily life [Kwak et al. 2010]. Performing sentiment analysis on tweets (Twitter microtexts) is not a trivial task due to the textual informality. Algorithms that carry out sentiment analysis on tweets usually adopt Natural Language Processing (NLP) techniques, such as POS Tagging. POS Tagging is used to detect subjective messages by identifying the grammatical classes of the words used in the text. POS Tagging tweets is not an easy task, since there are many abbreviations due to the character length constraint imposed by Twitter, being common to find repeated letters in many words as a way to emphasize terms, or even the absence of consonants in some words. Such problems get worse in texts written in Portuguese due to its grammatical complexity.

Many scientific proposals have arisen in the last years aiming to deal with sentiment analysis issues. However, most of them do not address both spatial and temporal dimensions enabling a most accurate analysis and a better understanding of the mood of people when using social media. Concerning the spatial dimension, the considered geographical location of a tweet can be the location where the social media user delivered the message, or the user's home location, or even a location eventually mentioned in the message. Concerning the temporal dimension, we are interested in possible opinion changes over time. Related work has considered the spatial dimension based only on geocoded messages. The main problem with such kind of approach is due to the fact that there are still few information sources which provide geocoded messages.

In this context, this work proposes a sentiment analysis approach which considers both spatial and temporal dimensions in order to better summarize the detected sentiment in a huge set of tweets harvested from the Web. Our approach rely on both Geographical Information Retrieval (GIR) and Sentiment Analysis techniques in order to infer geographical locations that may be mentioned and to detect the sentiment polarities from tweets, respectively. Combining spatial and temporal dimensions enables to visualize the sentiment spread over different geographical regions and temporal intervals.

We implemented and compared two approaches for sentiment analysis. We carried out a case study using tweets written in Portuguese related to the FIFA's Confederations Cup, which took place in Brazil in 2013. Both approaches address machine learning techniques for text classification, aiming to replace POS Taggers in the identification of opinionated tweets. The first approach is based on Naive Bayes classifiers whilst the second uses Support Vector Machines (SVM) classifiers. Each technique works with two classifiers: one to detect whether a tweet presents opinionated content and the other to classify the subjective polarity of the message as either positive or negative.

The main contributions of this article include the implementation and comparison of sentiment classification techniques for texts without using POS Taggers, which have shown good results when performed on Portuguese tweets; the application of GIR techniques in order to enable a spatial-driven analysis through the inferred geographical locations eventually mentioned in the text; and spatial and temporal sentiment analysis during the FIFA's Confederation Cup, held in Brazil in 2013. This work extends the work of Alves et al. [2014], including the spatial dimension to provide a sentiment map.

The remainder of this article is structured as follows. Section 2 highlights related work. Section 3 addresses the problem definition and challenges in the natural language processing in microblogs. Section 4 describes the case study into details. Section 5 discusses the achieved results. Finally, section 6 concludes the article and proposes further work to be undertaken.

2. RELATED WORK

Sentiment analysis has been used in many applications with several purposes [Liu 2012; Feldman 2013]: stock exchange companies, enabling the identification of the mood of the market based on specialists' opinions [Koppel and Shtrimberg 2006; O'Hare et al. 2009]; in the analysis of consumers' reviews about products or services [Eirinaki et al. 2012; Hu and Liu 2004]; the analysis of places or touristic regions by means of comments from tourists [Bjørkelund et al. 2012]; the analysis of politicians [Awadallah et al. 2012] or subjects related to politics [Fang et al. 2012]; and the real time monitoring of disease outbreaks in the regions of a country through sentiment analysis of messages posted on social networks [Silva et al. 2011].

Activities related to sentiment analysis comprise the detection of subjective or opinionated content, the classification of the content polarity and the summarization of the general sentiment of the evaluated entities. The sentiment detection in a text occurs in different levels: document, sentence and entity or aspects levels. Several methods have already been proposed to classify the sentiment polarity of a text and the main approaches used are based on machine learning techniques, semantic analysis techniques, statistical techniques and techniques based on lexical analysis or thesaurus. A comparison between these techniques can be found in [Sharma and Dey 2012] and an overview of sentiment analysis can be found in [Pang and Lee 2008] and [Liu 2012].

The sentiment analysis approaches that use machine learning implement classification algorithms such as Naive Bayes, Support Vector Machine (SVM), Maximum Entropy, Decision Trees (C4.5), K-Nearest Neighbor (KNN) and Condition Random Field (CRF). One of the main weaknesses on using supervised learning is the need of labeled data for training and evaluation tests. In order to help in the task of collecting the labeled data in an automated way, many proposals address the use of emoticons - characters used to transmit emotions. In the work of Li and Li [2011], 87% of the tweets which contain emoticons present the same sentiment represented in the text. Studies that employ emoticons to train the classifiers have presented excellent accuracy results (above 80%). The works of Go et al. [2009], Pak and Paroubek [2010] and Read [2005] report good results on sentiment classification using the Naive Bayes classifier.

Pak and Paroubek [2010] use the emoticons strategy to build the dataset to train a Naive Bayes classifier and categorize tweets as either positive or negative based on N-grams and in the grammatical classification of the words of the text by means of POS Tagger. One of the main problems in using only the emoticons in the collection of the data to train the classifiers is related to the recall metric, since emoticons are present in 10% of the tweets at least of the tweets [Gonçalves et al. 2013]. Our work differs from the work of Pak and Paroubek [2010] because it does not need POS Tagger to identify an opinionated (subjective) content. In addition, we have used manually annotated data to train the classifier besides the tweets with emoticons.

There are few works in the literature that perform sentiment analysis using a corpus in Portuguese. Chaves et al. [2012], Sarmento et al. [2009] and Tumitan and Becker [2013] use lexical analysis techniques based on thesaurus. Nascimento et al. [2009] apply machine learning techniques.

Chaves et al. [2012] present an algorithm called PIRPO (Polarity Recognizer in Portuguese) that uses a lexical analysis approach for sentiment classification from comments written in Portuguese. The algorithm uses ontologies and a list of polarized adjectives (positive, negative and neutral) that express sentiments to define the semantic orientation of the analyzed texts. The results achieved with PIRPO indicate an average of the F-Measure of only 0.32 at the recognition of polarity. Tumitan and Becker [2013] analyze the opinions from comments about politicians made in newspapers. They also study the correlation of expressed sentiments with the vote intention surveys. The polarity identification algorithm uses a word thesaurus (SentiLex-PT) that contains the polarity for every word (positive, negative, neutral), no matter the context.

Nascimento et al. [2009] use sentiment classifiers to evaluate the reactions of people on twitter concerning the news on the media. The results shown an accuracy varying from 70% to 80% according to the kind of news and the classifier applied.

Yu and Hatzivassiloglou [2003] use two classifiers. The first one is used to classify whether a tweet is either informative or opinative whilst the second one classifies the tweet polarity from the opinative tweets. A Naive Bayes classifier was used on both steps.

Faced with this scenario, our approach for sentiment analysis is similar to [Yu and Hatzivassiloglou 2003], which use two classification steps. However we propose the usage of a SVM classifier and compare it with a Naive Bayes classifier through two tasks: opinative tweet identification and polarity analysis. The adoption of two classifiers eliminates the need for using POS Tagger in the identification of an opinionated content. Thus, the first classifier detects whether the content is subjective or objective; and the second classifier identifies the polarity (positive or negative) of the content previously detected as opinionated. Furthermore, we apply GIR techniques in order to enrich the summarization, since it enables a spatial sentiment analysis.

3. THE SENTIMENT ANALYSIS APPROACH

Microblogging is a very popular communication tool among Internet users [Pak and Paroubek 2010]. The messages shared by microblogging users are not just about their personal lives, but also opinions and information about products, people, facts and events in general [Naaman et al. 2010]. The sites providing microblogging services, such as Twitter, become rich sources for mining user opinions. Since Twitter is a rich source of real-time information, many entities (i.e. companies, politicians, government, etc.) have demonstrated interest in knowing the opinions of people about services and products. The importance of Twitter for opinion mining has already been reported in other works [Kwak et al. 2010]. In this section, we describe in details our approach for spatial and temporal sentiment analysis on tweets.

3.1 A General Overview on Sentiment Analysis using the Spatial and Temporal Dimensions

The usage of Geographical Information Retrieval techniques, which is a subset of Information Retrieval, enables a summarization of the sentiment from a set of spontaneously exposed opinions in tweets. Such summarization encompasses both spatial and temporal aspects, enabling to visualize the sentiment variation over the time in specific geographical regions. Figure 1 shows the steps of the main flow for our sentiment analysis approach.

Once the tweets are harvested and stored in a local database, the pre-processing step starts. Since the text is essentially informal, many challenges must be taken into account in order to perform the sentiment analysis on tweets including grammatical errors, slang, and repeated characters. Hence, it is

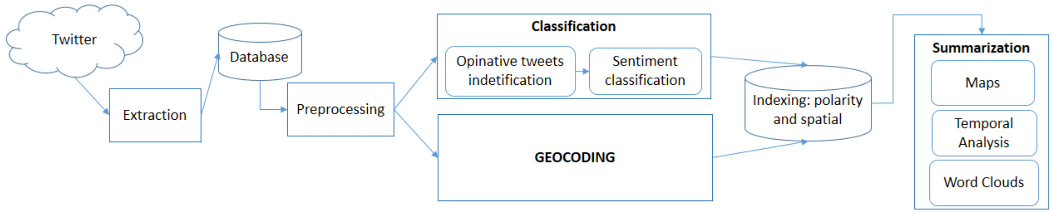


Fig. 1. Sentiment Analysis approach flow

required to deal with the text of a tweet in a specific manner. The literature presents some proposals for dealing with such information, namely: a) Filtering: removal of URLs, Twitter user names (starting with @) and Twitter special words ("RT", "via", ...); b) Removal of stop words; c) Use of synonyms for the decomposed terms; d) Part of speech tagging usage (POS tagging); e) Recognition/Extraction of entities; f) Stemming: method for reducing a term to its radical, removing the endings, affixes, and thematic vowels; and g) Treatment of the composite terms containing HashTags. The terms are normally separated according to the capitalization of the letters. For example, "#VeryGood" becomes "Very Good" - a blank space is added between the words.

The classification and geocoding steps are responsible for classifying the sentiment polarity and inferring geographical locations mentioned in the text, respectively. Both steps will be detailed in the following. Lastly, in the summarization step occurs the general sentiment modeling.

3.2 Classification of Sentiments Polarity

Sentiment analysis of tweets may be handled as a NLP task or, more specifically, as a text categorization task. Text categorization is the task defined as assigning predefined categories to text documents, in which documents can be news stories, technical reports, web pages, tweets, etc. These categories are most often subjects or topics, but may also be based on style (genres), pertinence, among others.

More formally, the text categorization task means finding a function that approximates the classification function $F : T \rightarrow C$, $f(t_i) = c_j$. This function describes how texts are associated to the classes, and also assigns a text $t_i \in T$ to its category $c_j \in C$, where T is a domain of texts and $C = \{c_1, ..., c_n\}$ is a set of n predefined categories.

In general, a text classification task starts with a training set $T = (t_1, ..., t_n)$ of texts that are already labeled with a category $c_j \in C$ (e.g. objective, subjective). The task tries to build a classification model (function F) that is able to assign the correct class to a new text t_i of the domain T.

In order to measure the performance of a classification model, a fraction of the labeled texts is set aside and not used for training. They are used to apply the proposed classification model and compare the estimated labels with true labels. To build the classification model (or to obtain the function F), we have implemented and evaluated two classifier techniques: SVM and Naive Bayes. In this work we implemented two approaches for sentiment classification of tweets and compared the achieved results. The sentiment classifying process for both approaches is performed in two classification steps (as shown in the diagram of Figure 1): in the first step only tweets with subjective text are classified; and in the last step each tweet is classified with a single sentiment either positive or negative. The SVM and Naive Bayes techniques have been used in the literature for traditional sentiment classification processes, in which several authors argue that SVM performs better in the text classification. Thus, this study aims at analyzing the behavior of such approaches when the sentiment classifying process is performed in two steps.

We also analyze the case study results through the standard Information Retrieval (IR) metrics, such as Precision, Recall and F-Measure. The fraction of correctly classified documents in relation to the total number of documents is called accuracy, and is a basic performance measure.

3.3 Detection of Geographical References

Some sentiment analysis proposals working on tweets consider the spatial dimension using only tweets already geocoded [Cho et al. 2014; Eshleman and Yang 2014]. Currently, the geographical location can only be found as metadata in a small amount of tweets. In addition to this issue, there are no guarantees that the geocoding information of those tweets refers to the geographical information mentioned in the text message. Other studies using tweets have shown that less than 5% of harvested tweets contain the geocoding information in its metadata [Alves et al. 2014]. Therefore, an alternative to increase the amount of georeferenced tweets is to take the usage of Geographical Information Retrieval (GIR) techniques in order to identify geographical locations in the text. Basically, GIR techniques try to identify place names into phrases and translate such place names into geographical coordinates. This is a challenging task that has some known open issues, since textual references for place names may be ambiguous, misspelled and too vague, leading to a wrong place name interpretation.

In order to infer the geographical locations in tweets, we have used the GeoSEn geoparser [Campelo and de Souza Baptista 2009] in this work. The GeoSEn geoparser is in charge of detecting geographic terms from texts written in Portuguese. It enables to infer toponyms eventually cited in a text following the Brazilian political-division hierarchy, which goes from the least precise levels (countries) to the most precise ones (cities). The process of detecting geographic references is based on a set of heuristics. All the details about how the GeoSEn geoparser works are described in [Campelo and de Souza Baptista 2009]. The choice of the GeoSEn system is due to its good achieved results for Web documents and also for tweets [de Oliveira et al. 2014; de Oliveira et al. 2015]. Other relevant factor we have considered on choosing the GeoSEn system is that it deals with Portuguese tweets and has an own gazetteer with Brazilian place names.

4. CASE STUDY DESIGN

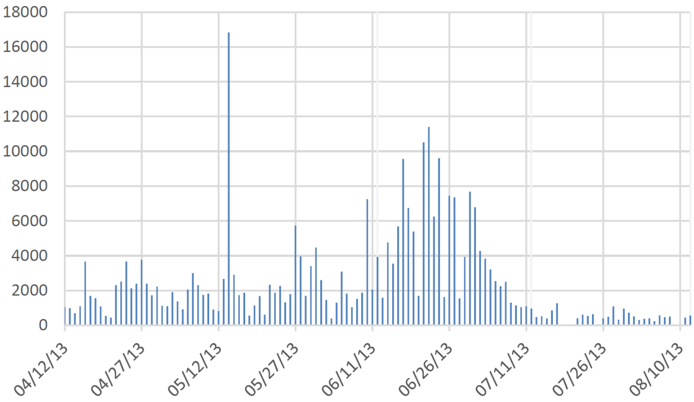
We present in this section the overall design of our developed case study. To describe this case study, this section is subdivided into two subsections: the selection of the corpus and the sentiment polarity classification of tweets.

4.1 Selection of the Corpus

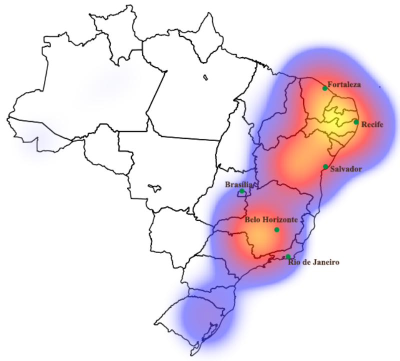
We collected approximately 300,000 tweets written in Portuguese regarding the FIFA's Confederations Cup, which took place in Brazil in 2013. The Twitter REST search API[1] was used for collecting tweets using GET requests. We implemented a crawler to harvest tweets automatically, every day between April and August 2013, which could contain at least one of the following terms: #copa2014, #Brasil2014, "Copa das Confederações" (Confederations Cup) and #copadasconfederacoes. Although our corpus is related to the 2013 Confederations Cup, we also have considered terms related to the 2014 World Cup since both competitions are quite related. The Confederations Cup is commonly considered a training competition for the World Cup in the following year. The tweets were stored in a database, and then submitted to a pre-processing step. Such step includes the removal of stop words, special terms (RT, via, etc.), removal of user names; and hashtags treatment (separation of composite terms, according to the capitalization of letters).

The data was collected between April 12th and August 12th (4 months) in the year of 2013, approximately two months before the beginning and two months after the end of the competition, which occurred from June 15th to June 30th, in 2013. We considered that period of time important for the temporal aspect of sentiment analysis regarding the opinion of the Brazilian people about the world cup in Brazil. Figure 2(a) illustrates the number of tweets posted every day during the time period the data has been collected.

[1] https://dev.twitter.com/rest/public/search



a)Number of collected tweets by day



b) Heat map from the geoparsed tweet set

Fig. 2. Collected Tweets

We have processed the entire tweet dataset using the GeoSEn geoparser on tweets as performed by de Oliveira et al. [2014]. Approximately 7,560 tweets were annotated with a geographical location. Figure 2(b) shows a heat map of the geographical distribution of those geoparsed tweets. Looking into the map we can notice a spatial co-relation between the geoparsed tweets and the regions of the FIFA's Confederations Cup 2013 host cities in Brazil.

It is possible to notice (Figure 2) that by the time of the competition (from June 15th to June 30th) the number of tweets was higher as expected. It is also possible to notice that, in May 14th of 2013, there were an unusual number of tweets. About 17,000 tweets were collected in that date, when the list of the summoned players from the Brazilian football team was announced in the press. Supposedly, most of those tweets were opinionated and must reflect the popular opinion about the selected players to form the team for the competition.

4.2 Sentiment Polarity Classification of the Tweets

To identify the opinative tweets and to classify their polarities we used Naive Bayes and SVM. Thus, we could compare the obtained results using these methods. In the approach of building a polarity classifier through supervised machine learning techniques, it was required to have labeled (classified) data to train the classifier. We then adopted two approaches to obtain the tweets with sentiments labeled:

Emoticons:. using the approach of Pak and Paroubek [2010], which associates the sentiment expressed by emoticons with all the words of the message which contains such emotional characters. These emotional characters could be happy emoticons - ":-)", ":)", "=)", ":D", etc - and sad emoticons - ":-(", ":(", "=(", ";(". Thus, if a tweet presents a happy emoticon (":-)"), for instance, its polarity is considered positive.

Manual Labeling:. 1,500 collected tweets were randomly chosen and separated for manual labeling of the sentiment polarity. We asked 10 volunteers to give their opinions about the sentiments present on these tweets. Only the tweets in which the volunteer opinions converged to the same sentiment are considered valid for labeling.

Both sentiment labeling methods were used to compare and combine the classifiers' results. Considering the randomness in the choice of tweets to be labeled manually, we found only 12 tweets with emoticons in which at least one of the volunteers confirmed the sentiments expressed by the emoticons. Table I presents the number of tweets with sentiments labeled using the two approaches. We considered neutral those tweets that do not express any opinion.

Table I. Number of Tweets Labeled (Training and Testing Sets)

Approach	Positive	Negative	Neutral	Total
Emoticons	1,468	492	-	1,960
Manual Labeling	326	321	463	1,110

We built two binary classifiers using distinct training sets: a classifier to check whether a tweet is subjective, which means it presents an opinion; and a polarity classifier to distinguish the sentiment as positive or negative. Once the sentiment classifier is trained, all tweets were analyzed and indexed with the opinion polarity obtained by the classifier. In order to validate the sentiment polarity classifier, we used 10-fold cross validation technique with all of the labeled tweets.

Finally, the collected sentiments were summarized regarding the temporal analysis. Such summarization has enabled users to browse through the overall mood of the sentiments expressed by the Brazilian people with respect to the subject of the Confederations Cup.

4.3 Summarization

Once the polarity detection and the geoparsing process on tweets have been finished, the achieved results were stored and indexed in a database in order to be handled by the summarization module. The spatiotemporal summarization comprises two subtasks, which are described as follows:

(1) Temporal Sentiment Analysis: associates both positive and negative sentiments along the time;
(2) Spatial Sentiment Visualization: provides sentiment heat maps which consider both tweet density and sentiment polarity in each geographical region. It enables to spatially summarize the sentiment.

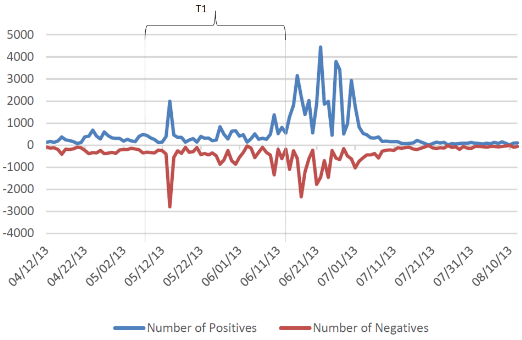
5. RESULTS AND DISCUSSION

Table II presents the results achieved for sentiment detection with the developed classifiers. The results obtained from the classifiers were summarized through the cross validation process using the entire annotated tweets dataset. The accuracy for the SVM classifier was 0.800 whilst the Naive Bayes classifier was 0.777, which makes the SVM classifier better. The F-Measure, a metric which correlates precision and recall, was also better using the SVM classifier. Thus, we can conclude the classifier which implements the SVM technique outperformed the Naive Bayes one presenting the best results.

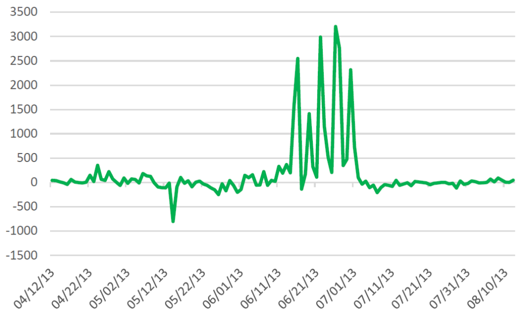
Once the sentiments classifier has been built, validated and also used all the labeled tweets, the next step was to obtain the general semantic orientation of the sentiments expressed by the Brazilian

Table II. Comparison of Developed Classifiers (SVM and Naive Bayes)

Classifier / Technique	**Dataset (Training and Testing)**	**Accuracy**	**Class**	**Precision**	**Recall**	**F-Measure**
Subjective tweet classification + Polarity Classifier / SVM	Emoticons + Manual Labeling	0.800	**Positive**	0.839	0.873	0.856
			Negative	0.715	0.657	0.685
			Weighted Average	0.799	0.802	0.800
Subjective tweet classification + Polarity Classifier / Naive Bayes	Emoticons + Manual Labeling	0.777	**Positive**	0.91	0.742	0.817
			Negative	0.616	0.849	0.714
			Weighted Average	0.813	0.777	0.783



a) Number of positive and negative tweets



b) Semantic orientation of the collected tweets

Fig. 3. Temporal Analysis of Polarity of Sentiment

people regarding the 2013 FIFA Confederations Cup. Figure 3(a) presents the result of the sentiments classifier applied to all collected tweets. In those charts, the number of sentiments classified with negative polarity was plotted in the negative semi-axis to avoid superposition with the number of sentiments classified with positive polarity.

One way of obtaining the general semantic orientation of the sentiment expressed in the tweets is subtracting the number of tweets with negative sentiments from the number of tweets with positive sentiments. Figure 3(b) presents the summary of the semantic orientation of the collected tweets.

Since we are interested in providing a spatial and temporal sentiment analysis, we consider a decision maker would be interested on understanding the sentiment distribution on both temporal dimensions - when the sentiment may vary as the time passes - and spatial dimension, where it can be possible to identify geographic regions with either negative or positive sentiment. In our approach, it is possible to select time periods from the data in order to visualize the heat maps. We have chosen the time period $T1$, shown in Figure 4, to exemplify such aspect.

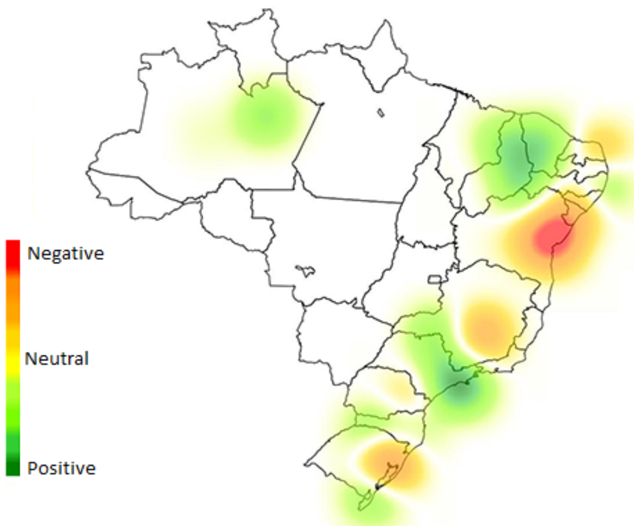


Fig. 4. Heat map from the sentiment-classified tweets of the time period $T1$

The $T1$, as shown in Figure 3(a), corresponds to a time period before the soccer competition starts. Two reasons led us to choose such time period: 1) it contains the date in which the Brazilian selected players were publicly announced, being a day which has registered too many tweets and also a good probability of most of them be opinative; 2) it is the period with the most negative sentiment tweets registered.

Figure 4 shows the spatial distribution of sentiment from tweets in the analyzed time period. It is possible to identify which Brazilian regions contributed to the increase of negative sentiment (represented in red in the figure), for example. When analyzing only the semantic orientation of the tweets (Figure 4) it is seen that the sentiment polarity throughout the time period was predominantly positive. On the other hand, when analyzing the spatial distribution of the sentiment, it is seen that the sentiment polarity varies through the regions.

Applying our approach for spatial and temporal sentiment analysis in the world business can enable decision makers to quickly interpose for reversing a negative sentiment about a product or a service offered in a specific geographical region. Thus, the sentiment analysis considering both spatial and temporal dimensions enables a more detailed analysis over an opinion set.

6. CONCLUSION

We have implemented and compared two approaches for sentiment analysis in a case study about tweets written in Portuguese related to the FIFA's Confederations Cup, occurred in Brazil in 2013. The first approach uses two Naive Bayes classifiers and the second one uses two SVM classifiers. The developed approaches replaced POS Tagger for identification of opinative tweets. The results obtained by the SVM sentiments classifier indicated an F-Measure of 0.873 and an accuracy of 80.0% for detection of the sentiment polarity. The Naive Bayes sentiment classifier presented an F-measure of 0.791 and an accuracy of 72.7%. The results presented by both classification approaches are considered satisfactory for Portuguese tweets, especially if we consider that the polarity of subjective content is not always consensual. The use of a SVM classifier increases the accuracy in 8%, which is a good result [Joachims 1998]. For example, in annotations made by humans, consensus hardly ever is above 75% [Pang and Lee 2008]. Other sentiment analysis studies applied to the English language have shown, at the best scenarios, accuracy around 95% for sentiment polarity detection [Sharma and Dey 2013]. The results achieved, however, are not enough to conclude whether it is always better to use SVM than Naive Bayes classifiers. It would be necessary novel studies with different datasets to reach a generalization.

We also performed a temporal analysis on the data, aiming to identify the semantic orientation of the sentiments expressed by means of the tweets, as well as the identification of the most cited terms in the opinionated messages. However, the main contribution of this work is the spatial and temporal summarization of the sentiment, which can be provided by using GIR techniques that enable inferring geographical locations in tweets. Our approach enables a more detailed sentiment analysis over an opinion set, since it provides both spatial and temporal views of the sentiment.

One of the main weaknesses identified is concerned with the identification of the entity referred to by the opinion detected in the tweet. Even though the collected tweets are related to the 2013 FIFA Confederations Cup, the opinions expressed in the messages may refer to other entities. In this sense, we believe that the application of Named Entity Recognition techniques might minimize this problem. Also, it will be interesting to investigate relationships across Twitter entities aiming at improving accuracy for detection of the sentiment polarity.

Further work will explore temporal series to help in the sentiment prediction according to a detected tendency. We also intend to analyze the opinion changes in order to provide maps with opinion trend flows according to the geographical location. Through the most frequent terms, we believe that it is possible to apply a similar approach used by Hu and Liu [2004], in which the automatic summarization of opinion about products and services reviews is performed and it includes aspects (features) from the observed entities. Thus, it will be possible to identify which aspects regarding the FIFA Confederations Cup were considered positive or negative also considering spatial and temporal dimensions.

REFERENCES

Alves, A. L. F., de Souza Baptista, C., Firmino, A. A., de Oliveira, M. G., and de Paiva, A. C. A Comparison of SVM Versus Naive-Bayes Techniques for Sentiment Analysis in Tweets: a case study with the 2013 FIFA Confederations Cup. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*. João Pessoa, Brazil, pp. 123–130, 2014.

Appelquist, D., Brickley, D., Carvahlo, M., Iannella, R., Passant, A., and Perey, C. A Standards-based, Open and Privacy-aware Social Web. http://www.w3.org/2005/Incubator/socialweb/XGR-socialweb-20101206/, 2010.

Awadallah, R., Ramanath, M., and Weikum, G. PolariCQ: polarity classification of political quotations. In *Proceedings of the International Conference on Information and Knowledge Engineering*. Sheraton, USA, pp. 1945–1949, 2012.

Bjørkelund, E., Burnett, T. H., and Nørvåg, K. A Study of Opinion Mining and Visualization of Hotel Reviews. In *Proceedings of the International Conference on Information Integration and Web-based Applications & Services*. Bali, Indonesia, pp. 229–238, 2012.

Calais Guerra, P. H., Veloso, A., Meira Jr., W., and Almeida, V. From Bias to Opinion: a transfer-learning approach to real-time sentiment analysis. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Diego, USA, pp. 150–158, 2011.

Campelo, C. E. C. and de Souza Baptista, C. A Model for Geographic Knowledge Extraction on Web Documents. In C. A. Heuser and G. Pernul (Eds.), *Advances in Conceptual Modeling - Challenging Perspectives*. Lecture Notes in Computer Science, vol. 5833. Springer, pp. 317–326, 2009.

Chaves, M. S., de Freitas, L. A., Souza, M., and Vieira, R. PIRPO: an algorithm to deal with polarity in Portuguese online reviews from the accommodation sector. In G. Bouma, A. Ittoo, E. Métais, and H. Wortmann (Eds.), *Natural Language Processing and Information Systems*. Lecture Notes in Computer Science, vol. 7337. Springer, pp. 296–301, 2012.

Cho, S. W., Cha, M. S., Kim, S. Y., Song, J. C., and Sohn, K.-A. Investigating Temporal and Spatial Trends of Brand Images Using Twitter Opinion Mining. In *Proceedings of the International Conference on Information Science and Applications*. Seoul, Korea, pp. 1–4, 2014.

de Oliveira, M. G., Baptista, C. S., Campelo, C. E. C., Acioli Filho, J. A. M., and Falcão, A. G. R. Automated Production of Volunteered Geographic Information from Social Media. In *Proceedings of the Brazilian Symposium on GeoInformatics*. Campos do Jordão, Brazil, pp. 118–129, 2014.

de Oliveira, M. G., Campelo, C. E. C., Baptista, C. S., and Bertolotto, M. Leveraging VGI for Gazetteer Enrichment: a case study for geoparsing twitter messages. In J. Gensel and M. Tomko (Eds.), *Web and Wireless Geographical Information Systems*. Lecture Notes in Computer Science, vol. 9080. Springer, pp. 20–36, 2015.

Eirinaki, M., Pisal, S., and Singh, J. Feature-based Opinion Mining and Ranking. *Journal of Computer and System Sciences* 78 (4): 1175–1184, 2012.

Eshleman, R. and Yang, H. "Hey #311, Come Clean My Street!" A Spatio-temporal Sentiment Analysis of Twitter Data and 311 Civil Complaints. In *Proceedings of the IEEE International Conference on Big Data and Cloud Computing*. Sydney, Australia, pp. 477–484, 2014.

Fang, Y., Si, L., Somasundaram, N., and Yu, Z. Mining Contrastive Opinions on Political Texts Using Cross-perspective Topic Model. In *Proceedings of the ACM International Conference on Web Search and Data Mining*. Seattle, USA, pp. 63–72, 2012.

Feldman, R. Techniques and Applications for Sentiment Analysis. *Communications of the ACM* 56 (4): 82–89, 2013.

Go, A., Huang, L., and Bhayani, R. Twitter Sentiment Classification Using Distant Supervision. http://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf, 2009.

Gonçalves, P., Araújo, M., Benevenuto, F., and Cha, M. Comparing and Combining Sentiment Analysis Methods. In *Proceedings of the ACM Conference on Online Social Networks*. Boston, USA, pp. 27–38, 2013.

Hu, M. and Liu, B. Mining and Summarizing Customer Reviews. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Seattle, USA, pp. 168–177, 2004.

Joachims, T. Text Categorization with Suport Vector Machines: learning with many relevant features. In C. Nedellec and C. Rouveirol (Eds.), *Machine Learning: ECML-98*. Lecture Notes in Artificial Intelligence, vol. 1398. Springer, pp. 137–142, 1998.

Koppel, M. and Shtrimberg, I. Good News or Bad News? Let the Market Decide. In J. G. Shanahan, Y. Qu, and J. Wiebe (Eds.), *Computing Attitude and Affect in Text: theory and applications*. The Information Retrieval Series, vol. 20. Springer, pp. 86–88, 2006.

Kwak, H., Lee, C., Park, H., and Moon, S. What is Twitter, a Social Network or a News Media? In *Proceedings of the International World Wide Web Conferences*. Raleigh, USA, pp. 591–600, 2010.

Li, Y.-M. and Li, T.-Y. Deriving Marketing Intelligence over Microblogs. In *Proceedings of the Hawaii International Conference on System Sciences*. Kauai, USA, pp. 1–10, 2011.

Liu, B. *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 2012.

Naaman, M., Boase, J., and Lai, C.-H. Is It Really About Me? Message Content in Social Awareness Streams. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. Savannah, USA, pp. 189–192, 2010.

Nascimento, P., Aguas, R., Lima, D. D., Kong, X., and Osiek, B. Análise de Sentimento de Tweets com Foco em Notícias. In *Proceedings of the Brazilian Workshop on Social Network Analysis and Mining*. Curitiba, Brazil, pp. 1–12, 2009.

O'Hare, N., Davy, M., Bermingham, A., Ferguson, P., Sheridan, P., Gurrin, C., and Smeaton, A. F. Topic-dependent Sentiment Analysis of Financial Blogs. In *Proceedings of the International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion Measurement*. Hongkong, China, pp. 9–16, 2009.

O'Reilly, T. What Is Web 2.0: design patterns and business models for the next generation of software. `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839`, 2007.

Pak, A. and Paroubek, P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *Proceedings of the International Conference on Language Resources and Evaluation*. Valletta, Malta, pp. 1320–1326, 2010.

Pang, B. and Lee, L. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2 (2): 1–135, 2008.

Read, J. Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification. In *Proceedings of the ACL Student Research Workshop*. Stroudsburg, USA, pp. 43–48, 2005.

Sarmento, L., Carvalho, P., Silva, M. J., and de Oliveira, E. Automatic Creation of a Reference Corpus for Political Opinion Mining in User-generated Content. In *Proceedings of the International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion Measurement*. Hongkong, China, pp. 29–36, 2009.

Sharma, A. and Dey, S. A Comparative Study of Feature Selection and Machine Learning Techniques for Sentiment Analysis. In *Proceedings of the Research in Applied Computation Symposium*. San Antonio, USA, pp. 1–7, 2012.

Sharma, A. and Dey, S. A Boosted SVM Based Sentiment Analysis Approach for Online Opinionated Text. In *Proceedings of the Research in Adaptive and Convergent Systems*. Montreal, Canada, pp. 28–34, 2013.

Silva, I. S., Gomide, J., Barbosa, G. A. R., Santos, W., Veloso, A., Meira Jr, W., and Ferreira, R. Observatório da Dengue: surveillance based on twitter sentiment stream analysis. In *Proceedings of the Brazilian Symposium on Databases, Demos Track*. Florianópolis, Brazil, pp. 49–54, 2011.

Tumitan, D. and Becker, K. Tracking Sentiment Evolution on User-generated Content: a case study on the Brazilian political scene. In *Proceedings of the Brazilian Symposium on Databases*. Recife, Brasil, pp. 1–6, 2013.

Yu, H. and Hatzivassiloglou, V. Towards Answering Opinion Questions: separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Stroudsburg, USA, pp. 129–136, 2003.