SINAI: Voting System for Aspect Based Sentiment Analysis

Salud María Jiménez-Zafra, Eugenio Martínez-Cámara,
M. Teresa Martín-Valdivia, L. Alfonso Ureña-López
SINAI Research Group
University of Jaén
E-23071, Jaén (Spain)
{sjzafra, emcamara, maite, laurena}@ujaen.es

Abstract

This paper describes the participation of the SINAI research group in Task 4 of the 2014 edition of the International Workshop SemEval. This task is concerned with Aspect Based Sentiment Analysis and its goal is to identify the aspects of given target entities and the sentiment expressed towards each aspect.

1 Introduction

The web has evolved progressively since its beginning in 1990. At first, the user was almost a passive subject who received the information or published it, without many possibilities to generate an interaction. The emergence of the Web 2.0 was a social revolution, because it offered users the possibility of producing and sharing contents, opinions, experiences, etc.

Some years ago it was common to ask family and friends to know their opinion about a particular topic, but after the emergence of the Web 2.0, the number of Internet users has been greatly increased. The exponential growth of the subjective information in the last years has created a great interest in the treatment of this information.

Opinion Mining (OM), also known as Sentiment Analysis (SA) is the discipline that focuses on the computational treatment of opinion, sentiment and subjectivity in texts (Pang and Lee, 2008). Currently, OM is a trendy task in the field of Natural Language Processing due mainly to the fact of the growing interest in the knowledge of the opinion of people from different sectors of the society. However, the study on Opinion Mining goes back to 2002 when two of the most cited articles in this task were published (Pang et al., 2002) (Turney, 2002).

OM or SA can be divided into two subtasks that are known as subjectivity classification and polarity classification. Subjectivity classification is the task concentrated on the identification of subjectivity in texts, that is, these systems are binary classifiers that separate the documents in two classes, objective and subjective ones. On the other hand, polarity classification is the task of determining the semantic orientation of a subjective text. The ideal OM system has to be composed by a subjectivity classifier and a polarity classifier. However, most of the works in the field of OM are carried out considering the documents as subjective, so polarity classification systems have been more studied than subjectivity classification ones. The reader can find a complete overview about the research in OM in (Pang and Lee, 2008) and (Liu, 2012).

As Liu asserts in (Liu, 2012), the polarity classification systems can be divided into three levels:

- **Document level polarity classification:** This kind of systems assumes that each document expresses an opinion on a single entity (Pang et al., 2002) (Turney, 2002).
- **Sentence level polarity classification:** In this case the polarity classification systems are focused on the identification of the level of polarity of each sentence of the document (Wilson et al., 2005) (Yu and Hatzivassiloglou, 2003).
- **Entity and Aspect level polarity classification:** These systems accomplish a finer-grained sentiment classification. Whereas the document-level and sentiment-level only discover the overall sentiment expressed by the author, the goal of the entity and aspect polarity classification is the identification of the

Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 566–571,
Dublin, Ireland, August 23-24, 2014.

sentiment of the author towards each entity or aspect.

An entity usually is composed by several aspects, for example a telephone is formed by a headset, which also consists of a speaker and an earphone. An entity can be regarded as a hierarchy of all the aspects whose head is the entity, so the entity can also be considered as an aspect or general aspect. Therefore, the task "entity and aspect level polarity classification" can be called "aspect polarity classification".

The main objective of OM at aspect level is to discover every quintuple (e_i, a_{ij}, s_{ijkl}, h_k, t_l) in a given document, where e_i is the entity, a_{ij} is one of the aspects of the entity or the entity and s_{ijkl} is the orientation of the opinion expressed by the opinion holder h_k in a certain moment t_l. To achieve the objective of populate the quintuple is needed the splitting of the task into several subtasks that correspond with the identification of the aspect, the author or the holder of the opinion and the moment when the opinion is expressed or posted. But in a real scenario, OM at aspect level is also limited like OM at sentence and document level, and most of the research works are only focused on the identification of the aspect and in the calculation of the level of intensity of the sentiment stated about the aspect. However, there are some papers that are closely to the goal of finding out each of the components of the quintuple (Kim and Hovy, 2004) (Kim and Hovy, 2006).

The task four of the 2014 edition of SemEval workshop aims to promote the research polarity classification systems at aspect level. The task is divided into four subtasks, two of them related to the aspect identification and the other with the polarity classification. Due to the fact that OM is a domain-dependent task, the organization proposes the four subtasks in two different domains, Restaurants and Laptops. Task one and three are the ones linked to the aspect identification. Subtask one is focused on the identification of the aspects in each review of the two given corpus. Subtask three goes one step further, in which the main objective is for a given predefined set of aspect categories, identify the aspect categories discussed in the given sentence. Subtask two proposes the classification of the sentiment expressed by the author about each of the aspects extracted, and subtask four has as challenge the classification of the polarity of each of the categories of the aspects. A wider description of the task and the datasets used can be found in the task description paper (Pontiki et al., 2014).

The rest of the paper is organized as follows. Section two outlines the two main parts of our proposed system, firstly the strategy to solve the subtask 1 and 2 and then the method used to resolve the subtask 3 and 4. To sum up the paper, an analysis of the results and the conclusion of this work are shown in section three and four respectively.

2 System description

The guidelines of this task indicate that each team may submit two runs: constrained (using only the provided training data and other resources, such as lexicons) and unconstrained (using additional data for training). We decided to follow an unsupervised approach that we present below.

Our system is divided into two subsystems (Figure 1). The aim of the first subsystem is to extract the aspect terms related to a given target entity (subtask 1) and calculate the sentiment expressed towards each aspect in the opinion (subtask 2). The goal of the second is, for a given set of categories, to identify the categories discussed in the review (subtask 3) and determine its polarity (subtask 4).

2.1 Subsystem 1: Aspects Identification and Polarity Classification

To identify the aspects related with the target entity (laptops or restaurants) we decided to use a bag of words built from all the aspect terms present in the training data. But this method only detects previously tagged aspect in the training data, so, we enriched the list of words with data automatically extracted from the collaborative knowledge base Freebase[1], in order to improve the identification. For this, we obtained all categories in restaurants domain and in computers domain[2] (types in a domain) using MQL[3] (Metaweb Query Language) (Figure 2).

Then, for each domain category we extracted all terms (instances of a type) to enrich the bag. In Figure 3 we can see an example to get all terms of

[1]http://www.freebase.com/

[2]Nowadays, Freebase has more than 70 different domains. But, for this task, we are only interested in these two.

[3]MQL is a language which is used to express Metaweb queries. This allows you to incorporate knowledge from the Freebase database into your own applications and websites.

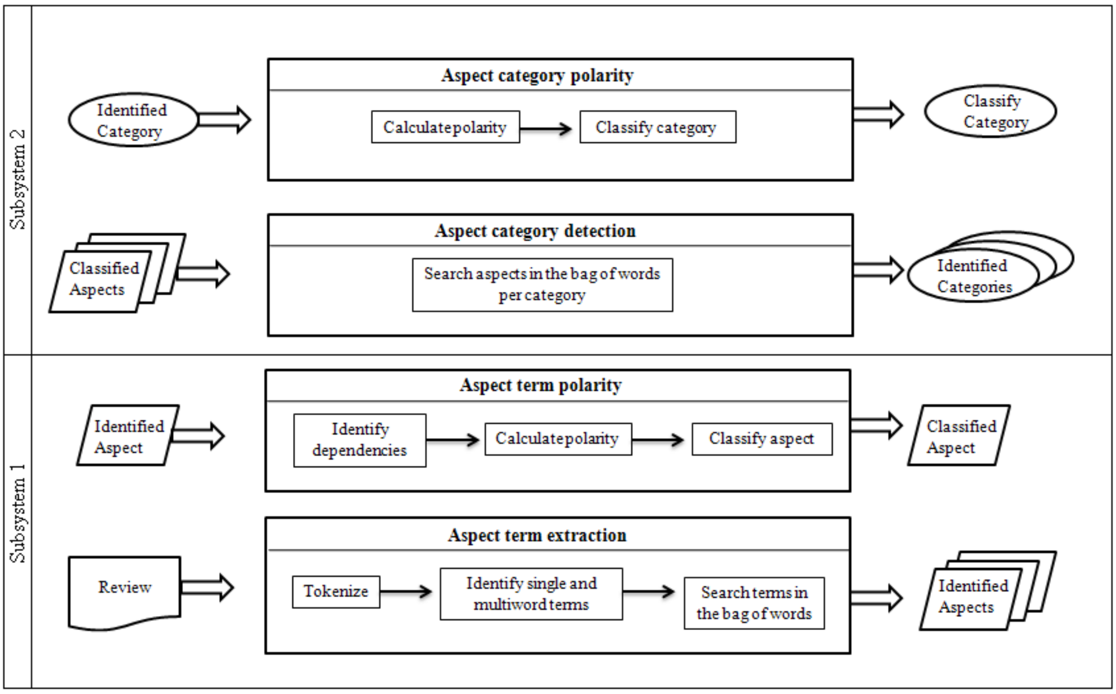

Figure 1: Arquitecture of the system.

```
[{
  "id": null,
  "name": null,
  "type": "/type/type",
  "domain": "/food"
}]
```

Figure 2: Query for list all categories in food domain.

a category, in particular cheese category of food domain.

```
[{
  "id": null,
  "name": null,
  "type": "/food/cheese"
}]
```

Figure 3: Query for list all term in cheese category.

In this way, given a review of the test data, the first step is to tokenize it to get a vector of unigrams with all single words in the text (we do not divide the reviews into sentences because there is only one sentence per review). The second step is to represent each review as a list of n lists of unigrams, bigrams, . . . , n-grams where n is the number of tokens in the sentence. This is because an aspect term can be a nominal phrase, a word formed from a verb but functioning as a different part of speech (e.g. gerunds and participles) or a simple term. For example, the review "The salad was excellent as was the lamb chettinad" is represented as shown in Figure 4.

After obtaining the possible terms of a review, the next step is to go over the list of lists to extract the aspects. Each list is traversed backwards matching each term with each aspect from the bag. When an aspect is found or the top of the list is reached the search begins in the next list. In the review showed in Figure 4, the system will identify two aspects: **salad** and **lamb chettinad**. The search in this example begins in the list 1 with "The salad was excellent as was the lamb chettinad", ends with "The" and continues with the next list, because the top of the list is reached. The search in the list 2 begins with "salad was excellent as was the lamb chettinad", ends with "salad" because it is an aspect and continues with the list 3 and so on. At last, the search in the list 8 begins with the term "lamb chettinad", ends with it because it is an aspect presents in the bag of words and continues with the list 9.

Once extracted the aspects related with the target entity, the next step is to determine the words that modify each aspect. For this, we have used the Stanford Dependencies Parser[4]. This parser

[4] http://nlp.stanford.edu/software/lex-parser.shtml

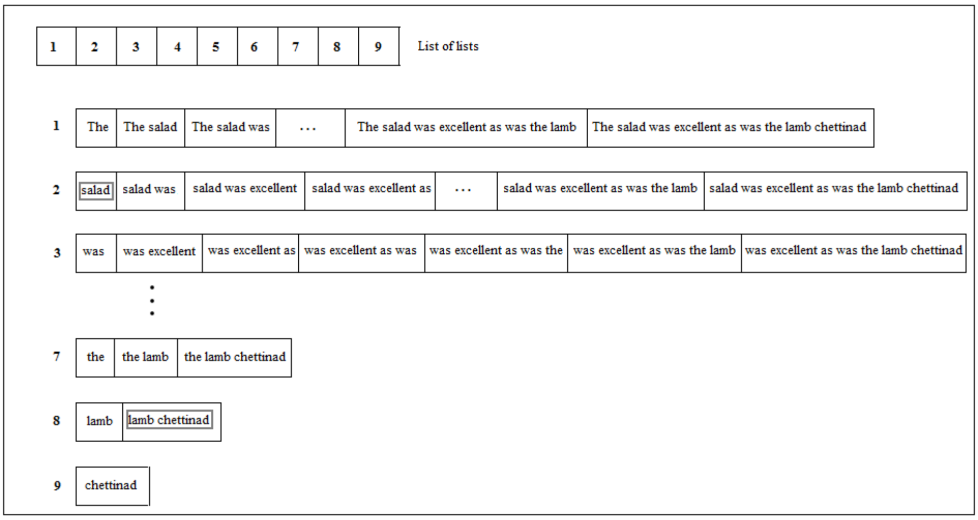

Figure 4: Possible terms of the sentence "The salad was excellent as was the lamb chettinad".

was designed to provide a simple description of the grammatical relationships in a sentence that can easily be understood and effectively used by people without linguistic expertise who want to extract textual relations (De Marneffe and Manning, 2008). It represents all sentence relationships uniformly as typed dependency relations. In this work, we have considered the main relationships for expressing opinion about an aspect: using a verb ("nsubj" or "nsubjpass"), an adjectival modifier ("amod") or a dependency relation with another word ("dep"). In the review "The salad was excellent as was the lamb chettinad", the system will identify two modifiers words: the adjective **excellent** that expresses how is the **salad** through the relationship **"nsubj"** and the adjective **excellent** that also modified the aspect **lamb chettinad** through the relationship **"dep"** Figure 5.

To determine the sentiment expressed over an aspect we have calculated the polarity of each word that modifies it through a voting system based on three classifiers: Bing Liu Lexicon (Hu and Liu, 2004), SentiWordNet (Baccianella et al., 2010) and MPQA (Wilson et al., 2005). The Bing Liu Lexicon is a list of 2006 positive words and another with 4783 negative ones. MPQA is also a subjectivity lexicon with positive and negative words and has extra information about each one: the part-of-speech, the strength, etc. Finally, SentiWordNet is a lexical resource that assigns to each synset of WordNet three sentiment scores: positivity, negativity and objectivity. Therefore, an aspect is positive/negative if there are at least two classifiers that tag it as positive/negative and neutral in another case. It may happen that a word is affected by negation, to treat this problem we have used a straightforward method, the fixed window size method. We have considered the negative particles: "not", "n't", "no", "never". So if any of the preceding or following 3 words to one aspect is one of these negative particles, the aspect polarity is reversed (positive —> negative, negative —> positive, neutral —> neutral).

In the example showed in Figure 5, the aspect salad is modified by the word excellent that also modified the aspect lamb chettinad. This adjective is part of the Bing Liu positive list, MPQA classifies it as positive and SentiWordNet assigns it the scores: 1 (positivity), 0(objectivity), 0 (objectivity). Then, the aspects **salad** and **lamb chettinad** are classified as **positive** by the voting system.

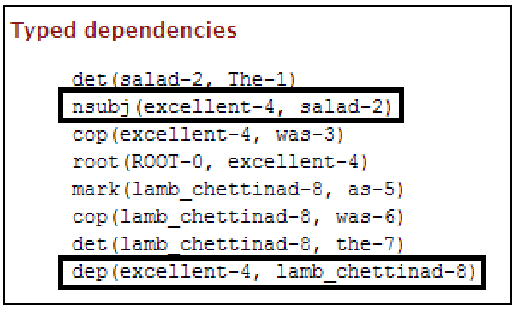

Figure 5: Dependency analysis of the sentence: "The salad was excellent as was the lamb chettinad".

2.2 Subsystem 2: Categories Identification and Polarity Classification

As we have mentioned above, this subsystem focuses on the treatment of the categories and has been used only with the dataset of restaurants.

On the one hand, we have built a bag of words for each of the given categories related to the target entity (restaurants). We have tagged manually each aspect of the bag of words, built for the first subsystem, in one of the categories of the given set (food, service, price, ambience, anecdotes/miscellaneous). Thus, to determine the categories that are referenced in a review we have searched each aspect identified with the first subsystem in each bag, if the aspect belongs to any category then this category is identified. If any aspect belongs to a category, then the category allocated is "anecdotes/miscellaneous".

On the other hand, the sentiment expressed about each category has been calculated as the most frequent polarity of the aspects that belongs to this category. In case of a tie between positive and negative values, the polarity value conflict is assigned to the category. If any aspect belongs to the category, then the polarity value of the review is assigned to the category.

In the above example, the aspects salad and lamb chettinad belong to food's bag of words, so that the system will identify that the **category food** is discussed in this review and will assign it the **polarity** value **positive**, because the sentiment expressed about the two aspects that belongs to this category is positive.

3 Analysis of the results

The aim of this section is to provide a meaningful report of the results obtained after participation in the task related to Aspect Based Sentiment Analysis (ABSA). Table 1 shows the evaluation results for the aspect extraction subtask. As we can see, the recall overcomes the mean value of results of participants in both domains (laptops and restaurants), that is, the system identifies quite aspects of the corpus. However, the precision is lower because the system identifies aspects that are not considered by the organization, due to the fact that our bag of words contains more aspects than the tagged by the organization.

The results reached in the aspect term extraction subtask are similar (Table 2). It should be taken into account that the system is a general-domain sentiment classifier, so it does not use specific knowledge for each of the domains. This fact can be shown in the results reached in the task of polarity classification for the two domains, which are similar. Therefore, this subtask could be improved by taking into account the domain and other relationships for expressing opinion about an aspect apart from that we have treated ("nsubj", "nsubjpass", "amod", "dep").

	Laptops		**Restaurants**	
	SINAI	**Average**	**SINAI**	**Average**
Precision	0.3729	0.6890	0.5961	0.7674
Recall	0.5765	0.5045	0.72487	0.6726
F-score	0.4529	0.5620	0.6542	0.7078

Table 1: Aspect Term Extraction results.

	Laptops		**Restaurants**	
	SINAI	**Average**	**SINAI**	**Average**
Accuracy	0.5872	0.5925	0.5873	0.6910

Table 2: Aspect Term Polarity results.

On the other hand, the results in the identification of the categories discussed in a review have been high (Table 3) and even overcome the average recall of the participating systems. At last, Table 4 shows the result evaluation of the aspect category polarity subtask that are slightly lower than the average. These tables show that is possible to reach good results using a simple approach as described in subsection 2.2.

	Restaurants	
	SINAI	**Average**
Precision	0.6659	0.76
Recall	0.8244	0.7226
F-score	0.7367	0.7379

Table 3: Aspect Category Detection results.

4 Conclusion and future works

In SA can be differentiated three levels of study of a text: document level, sentence level and aspect level. The document level analysis determines the overall sentiment expressed in a review, while the sentence level analysis specifies for each sentence of a text, whether express a positive, negative or neutral opinion. However, these two types of anal-

	Restaurants	
	SINAI	**Average**
Accuracy	0.6030	0.6951

Table 4: Aspect Category Polarity results.

ysis do not reach the level of detail that an user wants when searches for information about a product. The fact that the overall sentiment of a product is positive does not mean that the author has a positive opinion about all aspects of that product, or the fact that is negative does not involve that everything about the product is bad.

In addition, the large amount of sources and the high volume of texts with reviews, make difficult for the user to select information of interest. Therefore, it is necessary to develop classification systems at aspect level that help users to make decisions and, on the other hand, that show companies the opinion that consumers have about their products, in order to help them to decide what to keep, what to delete and what to improve.

In this paper we have presented our first approach for the Aspect Based Sentiment Analysis that has been developed for the task four of the 2014 edition of SemEval workshop. After analyzing the evaluation results we consider that is possible to introduce some improvements we are currently working: domain adaptation in the polarity calculation, consideration of other relationships to determine which words modify an aspect and treatment of negation (in the system proposed we have used the fixed window size method). Also, in a near future we will try to extrapolate it to Spanish reviews.

Acknowledgments

This work has been partially supported by a grant from the Fondo Europeo de Desarrollo Regional (FEDER), ATTOS project (TIN2012-38536-C03-0) from the Spanish Government, AORESCU project (P11-TIC-7684 MO) from the regional government of Junta de Andalucía and CEATIC-2013-01 project from the University of Jaén.

References

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204.

Marie-Catherine De Marneffe and Christopher D Manning. 2008. Stanford typed dependencies manual.

Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '04, pages 168–177, New York, NY, USA. ACM.

Soo-Min Kim and Eduard Hovy. 2004. Determining the sentiment of opinions. In *Proceedings of the 20th International Conference on Computational Linguistics*, COLING '04, Stroudsburg, PA, USA.

Soo-Min Kim and Eduard Hovy. 2006. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, SST '06, pages 1–8, Stroudsburg, PA, USA.

Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1–167.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, 2(1-2):1–135, January.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, EMNLP '02, pages 79–86, Stroudsburg, PA, USA.

Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval)*.

Peter D. Turney. 2002. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pages 417–424, Stroudsburg, PA, USA.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, HLT '05, pages 347–354, Stroudsburg, PA, USA.

Hong Yu and Vasileios Hatzivassiloglou. 2003. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, EMNLP '03, pages 129–136, Stroudsburg, PA, USA.