MODELO NO SUPERVISADO PARA ANÁLISIS DE SNTIMIENTO BASADO EN ASPECTOS IN ESPAÑOL.

UNSUPERVISED MODEL FOR ASPECT-BASED SENTIMENT ANALYSIS IN SPANISH

[1]Carlos Henríquez Miranda, [1]Freddy Briceño Díaz and [2]Dixon Salcedo Morillo [3]Albeiro Cortés Cabezas

[1]Autonomous University of Caribbean, Barranquilla, Colombia
[2]Department of Computer Science and Electronics, Coast University, Barranquilla, Colombia
[3]Department of Electronic Engineering, Surcolombiana University, Neiva, Colombia

ABSTRACT

This paper presents an unsupervised model for Aspect-Based Sentiment Analysis in Spanish language, which automatically extracts the aspects of opinion and determines its associated polarity. The model uses ontologies for the detection of explicit and implicit aspects, and machine learning without supervision to determine the polarity of a grammatical structure in Spanish. The unsupervised approach used, allows implementing a system quickly scalable to any language or domain. The experimental work was carried out using the dataset used in Semeval 2016 for Task 5 corresponding to Sentence-level ABSA. The implemented system obtained a 73.07 F1 value in the extraction of aspects and 84.8% accuracy in the sentiment classification. The system obtained the best results of all systems participating in the competition in the three issues mentioned above.

Keywords: aspect-based; ontology, sentiment analysis; unsupervised machine learning.

1. INTRODUCTION

Currently, a large amount of data produced worldwide is very attractive to different government, commercial and industrial sectors, but the extraction of information and its processing makes this process very complex manually.

In consequence, for more than a decade we have been working on systems that allow analyzing a large amount of data automatically, based on advances in disciplines such as natural language processing (NLP), data mining and cloud computing, among others (Sidorov, Faizliev, & Balash, 2018).

Within the NLP there is the Sentiment Analysis (SA), an area that seeks to analyze the opinions, sentiment, values, attitudes, and emotions of people towards entities such as products, services, organizations, individuals, problems, events, themes and their attributes (Liu, 2012). The SA has shown a great tendency of investigation in the last years, in its great majority in the English language (Vilares, Alonso, & Goméz-Rodríguez Carlos, 2013), and (Henríquez & Guzmán, 2017). However, recent contributions have been realized in other languages such as Spanish (Henríquez, Guzmán, & Salcedo, 2016), (Plaza-Del-Arco, Martín-Valdivia, María Jiménez-Zafra, Molina-González, & Martínez-Cámara, 2016) (Cruz, Troyano, Enriquez, & Ortega Universidad de Sevilla AvReina Mercedes, 2008) and, in french (Cadilhac, Benamara, & Aussenac-Gilles, 2010) and in Chinese (W. Zhang, Xu, & Wan, 2012), as well as other languages.

The great majority of SA tries to detect the overall polarity (positive or negative) of a paragraph or a complete text (Steinberger, Brychcín, & Konkol, 2014). Other approaches are at the sentence level, classifying the sentiment expressed in each sentence (Pang & Lee, 2008), or classifying it with relation to the specific characteristics of an entity found in each sentence (Liu, 2015).

The first two approaches are sometimes incomplete in the face of the reality of organizations that want to know in detail the behavior of a product (Henríquez, Plà, Hurtado, & Luna, 2017). In contrast, the AS at the level of aspects, aims to identify the properties (aspects) of a product or an entity, and determine the polarity of that entity.

The third approach is the Aspect-Based Sentiment Analysis (ABSA) and aims to identify the properties (aspects) of a product or entity and determine the polarity of that entity. An aspect is an attribute or component of an entity. For example, in the phrase, "*The sound quality of this phone is extraordinary*" the aspect is "*sound*", the entity is "*telephone*", and the associated sentiment is "*extraordinary*" that has "*positive*" polarity.

Within ABSA two types of aspects are distinguished, the explicit and the implicit. The first one directly denotes the objective of the opinion and the second also represents the objective of the opinion of a document but does not appear explicitly in the text (Liu, 2015).

This paper discusses the results of the implementation of a model that automatically extract the aspects (explicit and implicit) of an opinion, identify possible sentiment and determine its polarity shown. The model is based on ontologies and unsupervised machine learning and seeks to reduce human participation throughout the process.

The rest of the article is organized as follows. Section 2 deals with background and similar work. Section 3 describes the methodology used. Section 4 shows the experiments along with their results, and, in the last section, we present the conclusions.

2. BACKGROUND AND RELATED WORKS

In the literature, we found few references to Aspect-Based Sentiment Analysis in Spanish, even less on implicit aspects (Pontiki et al., 2016). Most are limited to applying the same techniques and methods used and tested for the English language (Henríquez & Guzmán, 2017).

For the extraction of aspects, there are different approaches shown in the literature. Those that use a predetermined list of aspects (Wang, Lu, & Zhai, 2010), those that rely on counting names and phrases to calculate their frequency within a document (W. Zhang et al., 2012) and those that take advantage of the relationships between sentiment and aspects (Qiu, Liu, Bu, & Chen, 2011). In addition, there are more advanced approaches based on supervised learning (Marcheggiani, Täckström, Esuli, & Sebastiani, 2014) and on probabilistic inference (Xianghua, Guo, Yanyan, & Zhiqiang, 2013).

From the previous approaches, the great majority does not take into account the concept or sense of the words that represent the aspects. These are considered simple "labels" that are not located in the context of the opinion or in the domain of the entity to which you are referring. In consideration of the above, the approach proposed here, considered the meaning of the aspects and uses semantic techniques based on ontologies, which have been successfully applied in natural language processing (NLP) tasks such as information extraction, disambiguation of the meaning of words, automatic summary of texts, among others (Henríquez & Guzmán, 2016).

Ontologies consist of formal and explicit specifications that represent the concepts of a given domain and its relationships, that is, they are an abstract model of a domain, where the concepts used are clearly defined (Studer, Benjamins, & Fensel, 1998). The literature shows how ontologies have been used for sentiment analysis in (Zhou & Chaovalit, 2008), (Lau, Raymond Y.K., Lai, Chapmann C.L., Ma, Jian, & Li, 2009), (Lizhen, Xinhui, & Hanshi, 2012), (Peñalver-Martinez et al., 2014), (Cadilhac et al., 2010) and (Kontopoulos, Berberidis, Dergiades, & Bassiliades, 2013). A comparison of how they were used is in (Henríquez & Guzmán, 2016).

To determine the polarity in an ABSA, usually two strategies are used, the machine learning-based and the lexical-based. The machine learning approach is based on the application of an algorithm that learns from a set of example data; on the other hand, the lexical-based strategy needs a lexicon of sentiment or word dictionaries with its polarity to be able to process them.

The machine learning approach is classified into supervised and unsupervised learning; the first depends on the existence of previously labeled training documents, that is, they already have polarity assigned, while the second does not need, or does not have, prior knowledge of data labeled with polarity for the sentiment analysis. Supervised learning predominates over unsupervised learning and tends to achieve better classification results, due to a large number of tagged training documents. However, it is sometimes difficult to have these documents labeled because a person should normally be used for this task, which means that it is more feasible to collect documents not automatically labeled, which are those used by the unsupervised approach (Medhat, Hassan, & Korashy, 2014).

Within the literature related to ABSA, there are works such as (De Freitas & Vieira, 2013), where the authors carried out an analysis supported by ontologies in the cinema and hotels domain in Portuguese. Thus, (Steinberger et al., 2014) presents a supervised approach in restaurant reviews in Czech, (Manek, Shenoy, & Mohan, 2016) proposes a system in the English language, based on the GINI index on cinema.

Recently unsupervised approaches have been used such as (Jiménez-Zafra, S. M., Martín-Valdivia, M. T., Martínez-Cámara, E., & Ureña-López, 2015) and (C. Wu, Wu, Wu, Yuan, & Huang, 2018) in English and (García-Pablos, Cuadros, & Rigau, 2018) in multiple languages. Another semi-supervised approach and special consideration are given in (Henríquez et al., 2017), a system in Spanish for ABSA that combines an unsupervised model for extraction of aspects and supervised machine learning for classification of sentiment.

3. METHODOLOGY

In Fig. 1, the proposed model is shown. This model consists of four layers: language processing, aspects extraction, sentiment identification and sentiment classification.

3.1 Layer 1: Language processing

This layer allows the entry of opinions by the user through a document written in natural language, in Spanish. Then, a common process is applied to most models of sentiment analysis. Subsequently, the best techniques tested in the literature for this task are used (Dey & Haque, 2008) and (Haddi, Liu, & Shi, 2013).

The input of the opinions is done as a simple grammatical structure (SGS) and passes through a series of processes that end with an output expressed in a set of words labeled and lemmatized S(W, P, L).

Table 1 shows an opinion and the result that this layer would yield, and each word the grammatical category and the motto are shown.

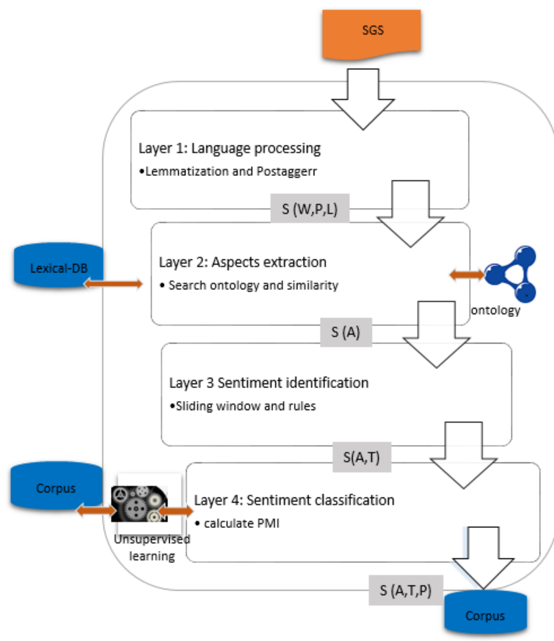

Fig. 1. Proposed model.

Table 1. Example of the process of lemmatization and postagger

Opinions	Results
OP2= {Quien sea amante de la carne tiene una carta bastante amplia para elegir, aunque ayer no tenían chuletón &&.}	S(W,P,L) = ("quien, P, quien", "sea, V,ser", "amante,N, amante", "de, S, de", "la, D, el", "carne, N, carne", "tiene, V, tener",

	“una, D, uno”, “carta, N, carta”, “bastante,R,bastante”, “amplia, A, amplio”, “para, S, para”, “elegir, V, elegir”, “aunque,C, aunque”, “ayer, R, ayer”, “ no, R, no”, “tenían, V, tener”, “chuletón,N,chuletón”, “.”, F)

3.1 Layer 2: Aspects extraction

To identify and extract the possible aspects of an entity from the opinions typed a semantic model MS is used (see Fig. 2). The model allows checking, if a set of the candidate aspects were found in the terminology of a specific domain with the help of a domain ontology and a lexical database.

The entry in this layer is a set of labeled and lemmatized words S(W, P, L), which is analyzed by the semantic model that determines a set of S(A) aspects identified as explicit and implicit.

Initially, the candidate aspects are taken (the word with grammatical category name) and a domain ontology is selected. The candidate aspects are compared with the classes and individuals of the ontology and those that match, are marked as explicit aspects.

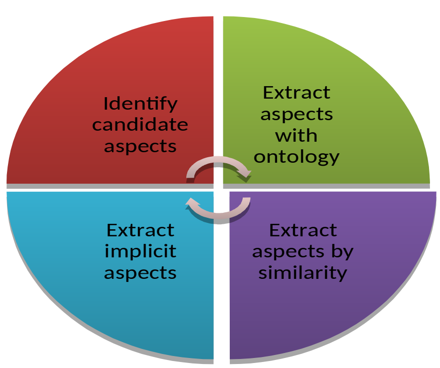

Fig. 2. Semantic model.

For example, if you have an ontology that models the domain of the hotels, (multilingual ontology "Hontology" of (Chaves, Larissa Freitas, & Renata Vieira., 2012)) and you have an opinion like "*Mi estancia en el hotel Dann fue gratificante. Las habitaciones estuvieron estupendas*", the semantic model can initially identify that "hotel" is an aspect since it coincided with an ontology class. “*Dann*" will be another aspect since it coincides with an individual. Finally, “*habitaciones*” will also be an aspect since it is a class related to “*hotel*” in Hontology.

After the previous process, the nouns of the opinions that were not found in the ontology undergo a process of semantic similarity with the ontology classes (Lan, Xu, & Gao, 2018). In this proposal, the calculation of the semantic similarity is based on the algorithm of Wu & Palmer (Z. Wu & Palmer, 1994) that considers the position of the concepts c1 and c2 in a taxonomy in relation to the position of the most specific common concept between the two (c1, c2), see equation 1.

$$¿_{\wp}(c1,c2)=\frac{2*depth(lso(c1,c2))}{len(c1,c2)+2*depth¿¿} \quad (1)$$

To find the similarity, the model considers that the *len* of the same concept is 0, *lso (c1, c2)* is the common ancestor, *depth (x)* is the depth from the root, and *depth (root) = 1*. For example, if you want to calculate the semantic similarity between two concepts like "*almuerzo*" and "*cena*" based on Palmer's distance, the taxonomy is shown in Fig. 3. Then, the depth from the root to the most common ancestor (*comida*) is equal to two (2), that is, *depth (lso ("almuerzo", "cena")* = 2, at the same time, if the length is 2, (len ("*almuerzo*", "*cena*") = 2), then *simwp ("almuerzo", "cena") = 0.667.*

To determine if a candidate for appearance is converted to an explicit aspect, the score of semantic similarity between the candidates and the concepts of the ontology is calculated, and then, is validated that the result is greater than or equal to an experimentally defined threshold.

For the extraction of implicit aspects in spanish, the best features obtained from the literature were combined with the use of domain ontology. In this study was used dual propagation techniques, which consist in a co-occurrence matrix between explicit aspects and opinion words to identify possible implicit aspects (W. Zhang et al., 2012), (Y. Zhang & Zhu, 2013) and (Sun, Li, Li, & Lv, 2014).

The implicit aspects are wanted in the opinions where there is no explicit aspect. To build the co-occurrence matrix, the double-propagation technique was used, starting with aspects candidates and the concepts of the first level of the domain ontology.

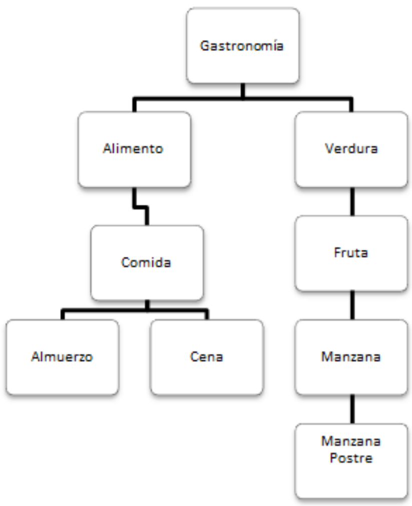

Fig. 3. An extract from the lexical database.

Finally, the output of this component of implicit aspects would be the aspects related to the implicit aspects found in the opinion. For example, if the opinion is "*no recomendable*", the component can throw for the opinion an explicit aspect related as "*comida*".

3.3 Layer 3: Sentiment identification

In this layer, the expressions were selected, based on their relationship with the aspects found in the previous layer, to find its polarity later. To achieve this, two techniques were used: sliding window and grammar rules.

The window process consists of taking the sentence where the aspect is and establishing a window of words to the right and left of the selected aspect. The default window length determined for this model was two (2) words. This value was defined experimentally for the restaurant domain.

With this length of the window, the purpose was to identify expressions of opinion that may affect the aspect. In the literature basically, adjectives (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011) have been used as expressions of opinion; this system had defined, based on the experimental phase, that expressions of opinion close to the aspect were adjectives and adverbs.

Additionally, grammar rules were used to determine if the sentiment found is affected by either negation or attenuation. Negation for his proposal is simple negation(Antònia Martí, Taulé, Teresa, Salud, & Jiménez-Zafra, 2016). The attenuation consisted on discovering the affectation of the sentiment by general adverbs like "*muy, bastante, demasiado, más* " among others.

Detecting either of these two situations alters the classification of sentiment in the next phase. The output of this phase will be a set of pairs, formed by aspect and expression of sentiment. Table 2 shows the possible output of the opinion: "*Las habitaciones grandes pero su mobiliario muy viejo. Se siente lúgubre. Las personas de la recepción muy amables. Piscina chévere* ".

Table 2. Example of aspects and expression of opinion

Aspect	**Expression of opinion**	**Grammatical feature**
Habitación	*Grandes, viejo*	
Mobiliario	*Grandes, viejo*	*Atenuación (muy)*
*	*Lúgubre*	
Recepción	*Amable*	*Atenuación (muy)*
Piscina	*chévere*	

You can notice an expression of opinion without the company of an aspect. In the example, the case of "*Lúgubre*" is shown, and is handled as an implicit aspect.

3.4 Layer 4: Sentiment classification

For the sentiment classification, a technique based on the measure of association, known as pointwise mutual information (PMI) (Church & Hanks, 1990), was used. This measure allows determining the semantic orientation of the expressions of opinion and the aspects through the appropriate selection of seeds of sentiment and a corpus of the domain.

The PMI of two words (x, y) is obtained by the probability that the two words appear together divided by the probabilities of each word individually (see equation 2).

$$PMI(x,y) = \log 2\left(\frac{P(x,y)}{P(x)P(y)}\right)^{\square} \qquad (2)$$

This was initially used by (Turney, 2002) in the sentiment analysis to calculate the semantic orientation of a sentence using the seeds "*excellent*" and "*poor*". Their idea was essentially to compare whether a phrase has a greater tendency to co-occur with the word "*poor*" or with the word "*excellent*" in a meta-search engine like *Altavista*.

In the proposed system, the calculation of the PMI was done for aspect-based sentiment analysis using: the aspect, the expression of opinion and a set of seeds.

To calculate the number of co-occurrences, the search engine is replaced by the count of occurrences and co-occurrences in a domain corpus formed by opinions without labeling. The PMI used considers only the positive values (Levy, Goldberg, & Dagan, 2015) and the irregular values that are presented are handled with an equilibrium factor.

For the calculation of the PMI, each expression of opinion x_i is taken and its frequency $f(x_i)_A$ is calculated only in the set of opinions in which aspect A appears. Then, we doing the same for each seed $f(y_j)_A$ and the co-occurrences between the two $f(x_i, y_j)_A$. With these values, a PMI greater than zero is obtained.

In the context of the proposed system, the $PMIP_A$ will be the highest PMI value between the expression of opinion and seed, see equation 3. Formally we have a set *n* of opinion expressions $\acute{X} = (S)$, and set of *m* seeds $\acute{Y} = (SE)$, and aspect *A*. Then, the positive pointwise mutual information $PMIP_A$ within a subset of opinions of the corpus where *A* is between $\acute{X} \wedge \acute{Y}$, will be the highest value between the concurrency of each seed y_j and the sentimental expression x_i.

$$PMIP_A(\acute{X},\acute{Y})\text{¿}\, max\left[\log 2\left(\frac{f(x_i,y_j)_A}{f(x_i)_A f(y_j)_A}\right)^{\square}\right]_{\square} \qquad (3)$$

In the previous calculation, there are cases in which the counting of the expression of opinion and the seed is rare along with aspect A. To address this case, the PMI calculated between the expression of opinion and each seed throughout the corpus and adjusts with the frequency of the expression of opinion as a balancing factor to prevent irregular values, and can see it in equation 4.

$$PMIP_A(\acute{X},\acute{Y})_{\square} = max\left[\log 2\left(\frac{f(x_i,y_j)_{\square}}{f(x_i)_{\square} f(y_j)_{\square}}\right)^{\square} * f(x_i)_{\square}\right]_{\square} \qquad (4)$$

The set of seeds defined for this work were five (5) words that represent an emotional disposition towards positive, negative and neutral. The words (seeds) selected for positive are “*excelente*” and “*bueno*, for negative “*malo*” and “*pésimo*” and for neutral “*indiferente* ".

Then, for the calculation of the polarity of each aspect, the PMI between each expression of opinion is calculated with the five seeds. From there you get the highest PMI value. If the highest PMI corresponds to the *"excelente"* and *"bueno"* seeds, the polarity is positive. If the highest PMI is from the *"malo"* and *“pésimo*" seed, the polarity is negative. Otherwise, the polarity would be neutral.

In the previous process, if an attenuation or a negation is found, the polarity given initially is changed. If attenuation is increased in one category (*bueno* instead of *excelente*) and if it is negation, the polarity is changed (*excelente* instead of *pésimo*). For the case of an implicit aspect, the nominal expression found with the related explicit aspect is taken for the calculation of PMI.

As the model receives more opinions, these are stored in the corpus of opinions adjusting the values of the PMI. The aspect, the opinion expression and the polarity will be stored in a database associated with the opinion.

The final output of the sentiment classification layer is a set of aspects with its expression of opinion and its associated polarity *S(A, T, P)* which is the final output of the model.

4. EXPERIMENTS AND RESULTS

The implementation of the model was done by building an application (*AspectSA*) under Java technology integrating different tools and libraries for the management of the Spanish language.

For the first phase of language processing, *Freeling* (Padró & Stanilovsky, 2012) was used for grammatical lemmatization and grammar labeling. For the aspects extraction, the ontology "*Hontology*" (Chaves et al., 2012) was used as a basis and adapted to the Spanish language (Fig. 4). Besides this, the same calculation of Multilingual Central Repository (*MCR*), that used as a *Wordnet* database in Spanish. For the sentiment classification, the corpus created in (Dubiau & Ale, 2013) was used, where 34808 positive and 16912 negative comments were obtained about restaurants. It is should be noted that the polarity of the corpus is not considered for validation.

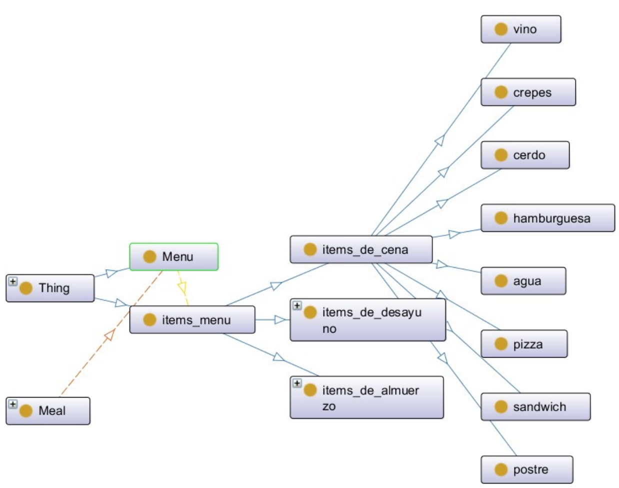

Fig. 4. An extract from the ontology "Hontology" used.

An example of the ABSA process performed by the *AspectSA* system shown in Table 3.

Table 3. An example of output AspectSA

Opinion	**Aspects Identified**	**Type**	**Sentiment**
Hotel viejo las habitaciones grandes pero su mobiliario muy viejo. Las personas de la recepción muy amables. Piscina chévere. Mala atención en el bar de la oficina. Pocas opciones de licor. No volveré.	*Hotel*	Explicit by ontology	Negative
	Habitación	Explicit by ontology	Positive
	Recepción	Explicit by ontology	Positive
	Piscina	Explicit by ontology	Positive
	Bar	Explicit by Similarity	Negative
	Licor	Explicit by Similarity	Negative
	**No volveré*	Implicit	Negative

In order to validate the proposed system, a series of experiments was carried out taking as reference the corpus of task 5 related to Aspect-Based Sentiment Analysis of the 2016 edition of *SemeEval* (International Workshop on Semantic Evaluation) an organization that performs, as a competence, continuous evaluations of computational systems of semantic analysis. Specifically, sub-task 1 (SB1) was addressed in the restaurant domain in spanish (Pontiki et al., 2016).

Subtask SB1, is divided into 3 subtasks, called slots. Slot 1 consists of detecting the category-aspect of an opinion. Each category is composed of an entity pair (E), attribute (A) represented by E # A.

Slot 2 consists of detecting the Opinion Target Expression (OTE) of a pair E # A, that is, the linguistic expression used in the opinion to refer to the entity (E) and the attribute (A). There may be opinions for which the OTE is null.

Finally, in the Slot 3 the polarity (positive, negative, neutral) of each OTE must be determined.

In the experiments of the proposed system, the following subtasks have been addressed: the subtask that deals with slot 1, slot 2 that corresponds to the aspects extraction in *AspectSA*, and slot 3 that corresponds to the sentiment classification in *AspectSA*. For this, the corpus (in Spanish) of the task consisting of 2070 training sentences and 881 evaluation sentences used. The metric evaluation for slot 2 F1 was used, and for slot 3, accuracy.

For the detection of the aspect category (slot 1), even the system is not suitable for this task, an adaptation was made taking advantage of the training data and the ontology. Table 4 shows the results of this task.

Table 4. SLOT 1 results

Variable	**Value**
Precision	55.06
Recall	66.41
F1	60.21

In Table 4, you can see that recall is higher than precision. Therefore, this indicates that the system for this domain correctly identifies many aspects and stops detecting only a few, however, the precision is lower because there were many false positives (many aspects were wrong). Likewise, F1 obtained a higher value, due to influence of recall.

Moreover, the extraction of aspects (subtask of slot 2), a set of experiments were performed and shown in Table V. In the original model, only the multilingual ontology "*Hontology*" is used, the adjusted model has the most important characteristics of the "*Restaurant*" (Recio-Garcia, 2006) ontology, semantic similarity is applied in the next row and in the last row it is applied additionally implicit aspects.

Table 5. Experiments for extraction of aspects (Slot 2)

Experiment	**F1**
Original model	61.9
Adjusted model	64.9
Adjusted model with similarity	65.58
Adjusted Model - with similarity and implicit aspects	73.07

In Table 5, it is can be seen that, as the number of features increases in the original model, the extraction of aspects increases. This increase is significant when all the characteristics of extraction, similarity and extraction of implicit aspects are completed, allowing the model to be more complete and robust.

For the classification of sentiment (polarity-Slot 3) we performed a series of experiments to adjust the unsupervised model, using the training corpus given by Semeval domain restaurants. The first experiment established what should be the expression of opinion appropriate to be related to the aspect and finally determine its polarity. Table VI shows the results of the experiments carried out that took into account: take all the expressions that accompany the aspect (column 2), take only adjectives (column 3), take adjectives and adverbs (column 4) and adjectives, adverbs and verbs (column 5).

Table 6. Experiment of expression of opinion

Variable/ Expression	**All**	**Adj.**	**Adj. y Adv.**	**Adj. Adv and Verb.**
Accuracy	54,55	83,53	83,61	74,83
Recall-positive	65,3	94,3	94,7	85,4
Recall-negative	28,8	54,7	55,2	56,4
Recall-neutral	15,6	5,8	5,1	2,7
Precision-positive	75,9	87,3	87,2	83,3
Precision-negative	33,9	64,9	66,7	54,2
Precision-neutral	4	27,3	27,3	6,3

Table 6 shows the behavior of each evaluation measurement for each of the expressions of opinion selected. You can observe the highest peak for accuracy is achieved when the expressions of opinion are adjectives and adverbs (83.61). It is can also be observed that the accuracy value of the system is due in large part to the high values of precision and positive completeness that the system throws.

With the best results (adjectives and adverbs) from the previous experiment, we set out to find the window length for the sliding window, which allows us to extract the opinion expressions appropriately. Table 7 shows the results of the experiments from a range of two (2) to ten (10) for the window length looking only for opinion expressions whose label is adverb or adjective.

Table 7. Experiment of the sliding window

Variable /length	2	3	4	5	6	7	8
Accuracy	85,47	84,53	83,65	82,31	81,58	80,81	80,22
Recall-positive	95,4	94,9	94,7	94,3	93,8	93,5	93,3
Recall-	54,6	55,5	55,2	53	53,5	52	50,2

negative							
Recall-neutral	6,8	5,8	5,1	4,8	3	2,9	4,3
Precision-positive	88,9	88,1	87,2	85,9	85,1	84,5	84
Precision-negative	63,4	66	66,7	65,1	65,4	63,8	62,8
Precision-neutral	50	27,3	27,3	30	25	25	33,3

According to the results shown in Tables 6 and 7 it could be established that the expressions of opinion to find the polarity of the aspects are adverbs and adjectives under a window length equal to 2 in the restaurant domain.

Then using the last configuration, we performed experiments to evaluation data the *Semeval*. For the slot 3 subtask, the corpus created in (Dubiau & Ale, 2013) was used, consisting of 34808 positive and 16912 negative comments about restaurants on the online food critic website www.guiaoleo.com. On this site, users express opinions about restaurants, and provide a rating in the category food, environment and service, assigning scores from 1 to 4 (bad/ regular, good, very good or excellent respectively).

Based on the corpus of (Dubiau & Ale, 2013), a balanced corpus was created with 40,000 opinions trying to have an equal number of positive and negative opinions. In addition, we not consider the general polarity of each opinion for the sentiment classification in the *AspectSA* system. After, we use this corpus to find the counting of the occurrences of each sentiment expression and seed, and their respective co-occurrences. Finally, we used in the process of double propagation and co-occurrence matrix in the extraction of implicit aspects.

In this experiment, we identified previously aspects used, and found the polarity from the opinion expressions. The sentiment classification results obtained from the *AspectSA* system are shown in Table 8.

Table 8. Slot 3 results

Variable	Value
Accuracy	84.8
Positive-Recall	94.1
Negative-Recall	53.1
Positive-Precision	89.1
Negative-Precision	50

Table 8 shows a high value in the accuracy influenced mostly by a recall and a high positive precision.

Table 9 shows the results of all the systems that participated in *Semeval* 2016, in the three categories described above, to establish a comparison with our system in the restaurant domain, subtask SB1 and Spanish language.

Table 9. Semeval 2016 results

Lang. /Dom. /Sub.	Slot 1 F1	Slot 2 F1	Slot 3 Accuracy
SP REST SB1	GTI/U/70.588 GTI/C/70.027 TGB/C/63.551 UWB/C/61.968 INSIG./C/61.37 IIT-T./U/59.899 IIT-T./C/59.062 UFAL/U/58.81 basel./C/54.686	GTI/C/68.515 GTI/U/68.387 IIT-T./U/64.338 TGB/C/55.764 basel./C/51.914	IIT-T./U/83.582 TGB/C/82.09 UWB/C/81.343 INSIG./C/79.571 basel./C/77.799

Tables IX show a list by the column of all the participants in the competition only in sub-task 1 (SB1), in the domain of restaurants (REST) and in the Spanish language (SP). In the list appears the name of the team followed by the letter U or C and then the value of the measure. The letter C indicates that it is restricted only to the training data provided and the letter U indicates unrestricted which allows the use of additional resources, such as lexical or training data. The table shows the values of measuring F1 for the first three tasks and the measure of accuracy for the last task. In the final part of each list, the baseline is shown as the initial reference value.

Table X shows the comparison between the results of the proposed system with the results of the winners of the *Semeval* competition.

TABLE X

Table 10. Comparison between Semeval and AspectSA

System Variable	Slot 1 F1	Slot 2 F1	Slot 3 Accuracy
As-	**60.21**	**73.07**	**84.8**
pectSA	70.58	68.515	N/A
GTI	59.899	64.338	83.582
IIT-T	63.551	55.764	82.09

TGB	61.968	N/A	81.343
UWB	61.37	N/A	79.571
INSIG	58.81	N/A	N/A
UFAL			

In Table 10, it is can be seen that *AspectSA* (proposed system) obtained the highest values in the extraction of aspects (slot 2) and the sentiment classification (slot 3) that all the proposed systems. This is highly significant, considering that an ABSA system must address the tasks of extracting aspects and sentiment classifying together. In detail it can be seen that the *GTI* (Alvarez-López, Juncal-Martinez, Fernández-Gavilanes, Costa-Montenegro, & González-Castano, 2016) system although it has high values of F1 in slot 1 did not obtain results for the sentiment classification. Additionally, the *IIT-T* (Kumar, Kohail, Kumar, Ekbal, & Biemann, 2016) system that has values similar to *AspectSA* in the sentiment classification (slot 3) is surpassed in the extraction of aspects (slot 2) by *AspectSA*, for more than 10 points. This allows inducing, in light of the results, that our *AspectSA* system is more robust and complete.

In Fig. 5, you can see the results of slot 2 of the *Semeval* competition and the *AspectSA* system. Here only four (4) teams participated with scores between 55.76 and 68.51 of F1. The *AspectSA* system greatly exceeded the best system of the competition by almost 5 points.

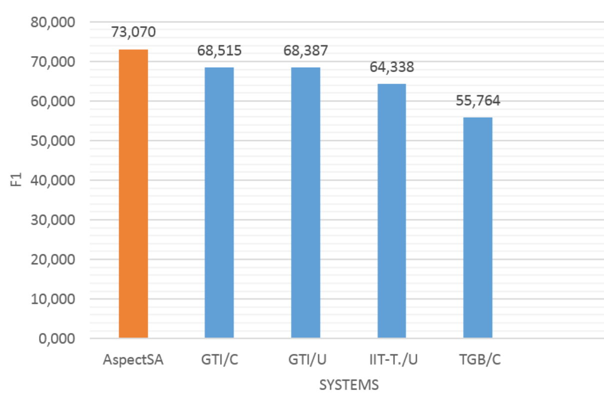

Fig. 5. Results of slot 2 of Semeval and AspectSA.

Analyzing the results of aspects extraction (slot 2), it should be noted that the choice and use of domain ontology, was vital for the identification of aspects since they represent the concepts of a given domain and its relationships. Additionally, these are an abstract model of a domain, of concepts clearly defined and are not in simple dictionaries. Moreover, reusing a validated domain ontology, in other tasks, it allowed an extraction that considered the meaning, because it was arranged in a specific domain already created, which allowed taking advantage of classes, individuals and relationships; In addition, it can exploit this knowledge of the domain, to improve the performance in the extraction of aspects.

The method of semantic similarity used in this work to address the extraction of aspects contributed significantly to the improvement of the process. For the evaluation set, an F1 value of 64.9 was improved, using only the ontology, to an F1 value of 73.07, obtained using ontology, semantic similarity and implicit aspects.

In Fig. 6, you can see the results of slot 3 of the *Semeval* competition and the system. Here only four (4) teams participated with scores between 79.57 and 83.58 of accuracy. Analyzing the polarity results, the proposed system achieved better results than those presented in *Semeval* obtained by the IIT-T team (Kumar et al., 2016). It should be noted that the proposed system works an unsupervised approach that does not depend on the domain and does not work with tagged data compared to (Kumar et al., 2016) that needs a tagged sentiment lexicon for the task. Additionally the IIT-T system has lower values in slot 1 and slot 2 than AspectSA, showing our system to be more complete for all the tasks of an ABSA.

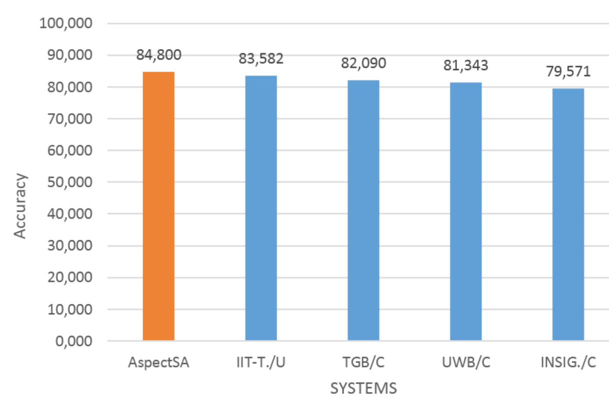

Fig. 6. Results of slot 3 of Semeval and AspectSA.

It is also important to indicate that the system only needs the domain corpus, the opinion expressions, and the seeds to obtain the sentiment associated. The larger the corpus, the easier the system can find more relationships between aspects with the words of opinion, but this causes a drop-in performance, so it was decided to work with the corpus of 40,000 opinions. Each time an opinion is processed, it is saved in the corpus, which allows the $PMIP_A$ calculated values to be adjusted little by little.

To prove this, an experiment was conducted in the domain of the hotels. For this experiment the original ontology "Hontology" the corpus "Hopinion" (http://clic.ub.edu/corpus/hopinion) in spanish was used, which contains about 17,934 opinions and 2,388,848 words, basically about hotels, from the TripAdvisor website. As there is no tagged corpus of evaluation for this task, 120 different opinions were taken from the Web in the domain of the hotels and were validated and analyzed by a human expert, who was in charge of determining the aspects of each opinion and its respective polarity. Each task was evaluated using 10-fold cross-validation. This option consists of dividing the data set into k equal and unique parts, that is, there cannot be the same sample in more than one part, and train the system with k-1 of the parts and verify it with the remaining part. This process is repeated k times, for each of the divisions of the data set. The results of the experiment are shown in Table 11.

Table 11. Results of experiments in the hotel domain

Measured	Value
Precision	91.66
Recall	86.84
F1	89.18
Accuracy	88.46

You can see in Table XI that the results thrown by the system were higher than the experiments carried out in the restaurant domain. This improvement shown can be explained from the fact that the data set has no spelling errors and most opinions do not have implicit aspects. These results were not compared with others because a common tagged corpus was not assigned for this task.

For the sentiment classification by *AspectSA*, the opinions of the corpus should be considered. If there is an unusual aspect in the domain, the system may throw out erroneous values or no value. This is compensated in part by finding $PMIP_A$ values only with the expressions of opinion surrounding the aspect and the seeds.

It is important to highlight the advantages of the system compared to the other systems that currently work for the Spanish language. The proposed system is one of the few existing systems that fully performs the process of aspect-based sentiment analysis in the Spanish language. In addition, is a completely unsupervised system that minimizes human presence for the two main processes of aspects extracting and sentiment classifying. This allows the system to be quickly scalable to any language or domain.

Moreover, we did in English language the experimentation with the *AspectSA* system. Therefore, the same multilingual ontology was used with the english part, and the corpus of opinions was changed by opinions in English. Finally, we compared the results of the experiments in *AspectSA* with the best *Semeval* result in English and are shown in Table 12.

Table 12. Results of experiments in English

Measured	AspectSA	Semeval
F1	60.86	72.34 basel./C/44.071
Accuracy	72.08	88.12 basel./C/76.484

In Table 12, you can see that the results of the system are more than acceptable, although we do not obtained the best results in the English language. Regarding, in baseline was exceeded the extraction of aspects, and sentiment, classification there is a difference of four (4) points; this shows that the system can be easily scaled to other languages with small changes.

5. CONCLUSIONS

Sentiment analysis (SA) has been the subject of research in recent years, due to the large-scale production of opinions by users on the Internet. However, the efforts had concentrated on performing an SA at the document level, which leads to not meeting the expectations of companies interested in knowing in detail the opinions and comments regarding their object of service or product.

For this reason, the aspect-based sentiment analysis (ABSA) has kept the attention of researchers, since it allows a fine-grained analysis very useful for different organizations and companies. This consists of two important tasks, the extraction of aspects and the sentiment classification of those aspects. However, not all systems address the two tasks with equal efficiency.

Likewise, the contributions in ABSA in spanish language are very few at the moment, for this reason, this study was aimed at building a system in spanish language, that would reduce human participation and achieve results comparable with existing systems.

The proposed system integrates ontologies and unsupervised machine learning and does not depend on the domain or tagged data and can be implemented in different languages with small changes.

The AspectSA system obtained a 73.07 F1 value in the aspects extraction and 84.8% accuracy in the sentiment classification. The system obtained the best results of all systems participating in the competition in the two aspects mentioned above. It is should be noted that the system addresses the tasks of an ABSA system with excellent results, showing a more robust and complete system compared to the systems participating in Semeval.

In particular, we can show different aspects that would improve the proposed system. Consequently, new research projects can be formulated that can give continuity to this work. Following are the main lines that could be developed:

(i) Build a tool that allows the assessment of the aspects extraction of aspects and sentiment classification by a human. (ii) Develop a tool that allows a human to review the explicit aspects by similarity and can decide if is added to the domain ontology as a concept or an individual. (iii) Explore new mechanisms that allow integrating the information and relationships of ontologies in automatic learning algorithms and be able to cover the tasks related to the sentiment analysis at the level of aspects. In the same way, it is could be extended to other languages and domains.

REFERENCES

Alvarez-López, T., Juncal-Martinez, J., Fernández-Gavilanes, M., Costa-Montenegro, E., & González-Castano, F. J. (2016). Gti at semeval-2016 task 5: Svm and crf for aspect detection and unsupervised aspect-based sentiment analysis. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)* (pp. 306–311).

Antònia Martí, M., Taulé, M., Teresa, M., Salud, M.-V., & Jiménez-Zafra, M. (2016). La negación en español: análisis y tipología de patrones de negación * Negation in Spanish: analysis and typology of negation patterns. *Procesamiento Del Lenguaje Natural*, (57), 41–48.

Cadilhac, A., Benamara, F., & Aussenac-Gilles, N. (2010). Ontolexical resources for feature based opinion mining : a case-study, 77–86.

Chaves, M., Larissa Freitas, & Renata Vieira. (2012). Hontology: a multilingual ontology for the accommodation sector in the tourism industry. In *CTIC/STI - Comunicações a Conferências*. Retrieved from http://hdl.handle.net/10884/654

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, *16*(1), 22–29.

Cruz, F. L., Troyano, J. A., Enriquez, F., & Ortega Universidad de Sevilla AvReina Mercedes, J. (2008). Clasificación de documentos basada en la opinión: experimentos con un corpus de críticas de cine en español Experiments in sentiment classification of movie reviews in Spanish. *Procesamiento de Lenguaje Natural*, *41*, 73–80.

De Freitas, L. A., & Vieira, R. (2013). Ontology-based Feature Level Opinion Mining for Portuguese Reviews. In *Proceedings of the 22nd International Conference on World Wide Web. ACM,* (pp. 367–370).

Dey, L., & Haque, S. M. (2008). Opinion mining from noisy text data. In *Proceedings of the second workshop on Analytics for noisy unstructured text data*.

Dubiau, L., & Ale, J. M. (2013). Análisis de Sentimientos sobre un Corpus en Español: Experimentación con un Caso de Estudio. In *Argentine Symposium on Arti_cial Intelligence,* (pp. 1850–2784).

García-Pablos, A., Cuadros, M., & Rigau, G. (2018). W2vlda: almost unsupervised system for aspect based sentiment analysis. *Expert Systems with Applications*, *91*, 127–137.

Haddi, E., Liu, X., & Shi, Y. (2013). The role of text preprocessing in sentiment analysis. In *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2013.05.005

Henríquez, C., & Guzmán, J. (2016). Las ontologías para la detección automática de aspectos en el análisis de sentimientos. *Revista Prospectiva*, *14*(2), 90–98. https://doi.org/10.15665/rp.v14i2.750

Henríquez, C., & Guzmán, J. (2017). A Review of Sentiment Analysis in Spanish. *Tecciencia*, *12*(22), 35–48. https://doi.org/10.18180/tecciencia.2017.22.5

Henríquez, C., Guzmán, J., & Salcedo, D. (2016). Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles Opinion. *Procesamiento Del Lenguaje Natural*, *41*, 25–32.

Henríquez, C., Plà, F., Hurtado, L. F., & Luna, J. A. G. (2017). Análisis de sentimientos a nivel de aspecto usando ontologías y aprendizaje automático. *Procesamiento Del Lenguaje Natural*, *59*, 49–56.

Jiménez-Zafra, S. M., Martín-Valdivia, M. T., Martínez-Cámara, E., & Ureña-López, L. A. (2015). Combining resources to improve unsupervised sentiment analysis at aspect-level. *Journal of Information Science*.

Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2013.01.001

Kumar, A., Kohail, S., Kumar, A., Ekbal, A., & Biemann, C. (2016). IIT-TUDA at SemEval-2016 Task 5: Beyond Sentiment Lexicon: Combining Domain Dependency and Distributional Semantics Features for Aspect Based Sentiment Analysis. In *Proceedings of SemEval* (pp. 1129–1135).

Lan, M., Xu, J., & Gao, W. (2018). Ontology similarity computation and ontology mapping using distance matrix learning approach. *IAENG International Journal of Computer Science*, *45*, 164–176.

Lau, Raymond Y.K., Lai, Chapmann C.L., Ma, Jian, & Li, Y. (2009). Automatic domain ontology extraction for context-sensitive opinion mining (pp. 35–53).

Levy, O., Goldberg, Y., & Dagan, I. (2015). Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, *3*, 211–225.

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Sentiment Analysis and Opinion Mining*.

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. (Cambridge University Press, Ed.).

Lizhen, L., Xinhui, N., & Hanshi, W. (2012). Toward a fuzzy domain sentiment ontology tree for sentiment analysis. In *Image and Signal Processing (CISP), 2012 5th International Congress on. IEEE*.

Manek, A. S., Shenoy, P. D., & Mohan, M. C. (2016). Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World Wide Web*, 1–20.

Marcheggiani, D., Täckström, O., Esuli, A., & Sebastiani, F. (2014). Hierarchical multi-label conditional random fields for aspect-oriented opinion mining. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-06028-6_23

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*. https://doi.org/10.1016/j.asej.2014.04.011

Padró, L., & Stanilovsky, E. (2012). FreeLing 3.0: Towards Wider Multilinguality. In *LREC2012*.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(12), 1–135.

Peñalver-Martinez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Ángel Rodríguez-García, M., Moreno, V., Fraga, A., & Sánchez-Cervantes, J. L. (2014). Feature-based opinion mining through ontologies. *Expert Systems with Applications* , *41*(13), 5995–6008. https://doi.org/10.1016/j.eswa.2014.03.022

Plaza-Del-Arco, F. M., Martín-Valdivia, M. T., María Jiménez-Zafra, S., Molina-González, M. D., & Martínez-Cámara, E. (2016). COPOS: Corpus Of

Patient Opinions in Spanish. Application of Sentiment Analysis Techniques. *Procesamiento Del Lenguaje Natural*, *57*, 83–90.

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., … Eryiğit, G. (2016). SemEval-2016 Task 5: Aspect Based Sentiment Analysis. In *Semeval* (pp. 19–30).

Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion Word Expansion and Target Extraction through Double Propagation. *Computational Linguistics*, *37*, 9–27.

Recio-Garcia, J. (2006). restaurant.owl. Retrieved February 1, 2017, from http://www.disi.unige.it/person/LocoroA/download/wilfontologies/restaurant.owl

Sidorov, S., Faizliev, A., & Balash, V. (2018). Fractality and multifractality analysis of news sentiments time series. *IAENG International Journal of Applied Mathematics*, *48*, 90–97.

Steinberger, J., Brychcín, T., & Konkol, M. (2014). Aspect-Level Sentiment Analysis in Czech. In *Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 24–30).

Studer, R., Benjamins, V. R., & Fensel, D. (1998). I DATA & KNOWLEDGE ENGINEERING. *Data & Knowledge Engineering*, *25*, 161–197.

Sun, L., Li, S., Li, J., & Lv, J. (2014). A novel context-based implicit feature extracting method. In *Data Science and Advanced Analytics (DSAA), 2014 International Conference on* (pp. 420–424).

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics* , *37*(2), 267–307.

Turney, P. D. (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. Association for Computational Linguistics* (pp. 417–424).

Vilares, D., Alonso, M. a., & Goméz-Rodríguez Carlos. (2013). A syntactic approach for opinion mining on Spanish reviews. *Natural Language Engineering*, *1*(1), 1–26.

Wang, H., Lu, Y., & Zhai, C. (2010). Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach. ACm.

Wu, C., Wu, F., Wu, S., Yuan, Z., & Huang, Y. (2018). A hybrid unsupervised method for aspect term and opinion target extraction. *Knowledge-Based Systems*, *148*, 66–73. https://doi.org/10.1016/J.KNOSYS.2018.01.019

Wu, Z., & Palmer, M. (1994). Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics* (pp. 133–138).

Xianghua, F., Guo, L., Yanyan, G., & Zhiqiang, W. (2013). Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. *Knowledge-Based Systems*. https://doi.org/10.1016/j.knosys.2012.08.003

Zhang, W., Xu, H., & Wan, W. (2012). Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2012.02.166

Zhang, Y., & Zhu, W. (2013). Extracting Implicit Features in Online Customer Reviews for Opinion Mining. In *Proceedings of the 22Nd International Conference on World Wide Web* (pp. 103–104). New York, NY, USA: ACM. https://doi.org/10.1145/2487788.2487835

Zhou, L., & Chaovalit, P. (2008). Ontology-supported polarity mining. *Journal of the American Society for Information Science and Technology*, *98–110*.