Aspect Extraction for Opinion Mining with a Semantic Model

Carlos Henriquez and German Sanchez-Torres

***Abstract*—In this article, we present a semantic model for aspect extraction from Spanish text as part of a complete aspect-based sentiment analysis system. The model uses ontology, semantic similarity, and double propagation techniques to detect explicit and implicit aspects. The proposed approach allows the implementation of a scalable system for any language or domain. The experimental tests were carried out using the *SemEval-2016* dataset for task 5, corresponding to the aspect-based sentiment analysis sentence level. The implemented system obtained an F1 score of 73.07 for the aspect extraction, achieving the best results among the systems participating in the comparison, and an F1 score of 89.18 for the hotel domain using a ten-iteration cross-validation.**

***Index Terms*—Aspect-based sentiment analysis, ontology, opinion mining, natural language processing, semantic similarity.**

I. Introduction

Currently, the amount of data produced worldwide is very large due to the massive use of social networks, messaging services, blogs, wikis, and e-commerce, among others. This range of data is attractive in commercial, industrial, and academic scenarios, among others. However, this makes the task of extraction and its respective manual processing very complex and difficult to perform [1]. Because of the potential use of analyzing and managing these data, there are large work efforts to find models, techniques, and tools that allow automatic text analysis [2]. Recent research has delved into a type of natural language processing (NLP) called Sentiment Analysis (SA) or Opinion Mining (OM). OM seeks to analyze the opinions, sentiments, evaluations, attitudes, and emotions of people towards products, services, organizations, individuals, problems, events, and problems as well as their characteristics [3].

According to [4], in sentiment analysis, there are three levels: the document, sentence, and aspect levels. At the document level, the whole-document sentiment can be classified as positive or negative [5]. At sentence level, the purpose is to classify each sentence as positive, negative, or neutral, and finally, at aspect level, the aim is to perform a classification with respect to the specific characteristics of each of the entities [6].

At present, the vast majority of approaches to OM detect sentiments at a general level in a complete sentence or document [7]. However, these approaches are incomplete with respect to the reality of companies that seek to know in detail what is written about their product [8]. According to [3], *the document level and the phrase level do not discover what exactly people like and dislike, unlike the OM at the aspect level, which performs the analysis in more detail*; that is, it focuses on the fundamental characteristics of the opinion. OM at the level of aspects aims to identify the properties of an entity and the sentiments associated with an expression. That is, in a written text that represents an opinion, the fundamental characteristics, or aspects, of an entity must be identified and then an associated entity determines the polarity, which can be positive or negative. The entity represents the part being reviewed or commented on and can be a person, product, company, or any type of object or reason for commenting. An example of OM at the level of aspects can be demonstrated for the sentence ”La calidad del sonido de este teléfono es increíble”; here the aspect is ”sonido”, the entity is ”teléfono”, and the associated sentiment is ”increíble”, which has ”positive” polarity.

In this OM approach, two types of aspects are distinguished. The first refers to the explicit aspects, which are words in the document that directly denote the objective of the opinion. The second is the implicit aspect, which represents the objective of the opinion of a document but is not explicitly specified in the text [9]. This article shows the results of the implementation of a model that makes it possible to automatically extract the aspects (explicit and implicit) of a text for an opinion mining system. The system is based on a semantic model that integrates ontologies, semantic similarity, and double propagation techniques together with a co-occurrence matrix. The rest of the article is organized as follows. Section II addresses the background and similar work, Section III describes the methodology used, Section IV presents the experiments and results, and finally, the conclusions are given in Section V.

II. Background

Aspect-level OM, also known as aspect-based sentiment analysis or feature-based AS, identifies the properties or characteristics of an entity and determines the expressed polarity of every aspect of that entity [10]. According to [6], there are two tasks related to OM at the aspect level. The first task is related to detecting and extracting aspects of an entity in a given text and the second task is to determine the sentiment associated with that aspect or its polarity.

Different approaches have been used for the first task. Below is a compilation of some important papers that address this task through different approaches [4].

The first approach found for the extraction of aspects is based on counting names and phrases to calculate their frequency within a document [11], [12]. Another approach used is to take advantage of the relationships of aspects

Manuscript received April 07, 2020; revised June 03, 2020.

This work was supported by the Universidad del Magdalena, Santa Marta, Colombia.

C. Henriquez is with Systems and Computing Research Group, Faculty of Engineering, Universidad del Magdalena, Santa Marta, Colombia e-mail: chenriquezm@unimagdalena.edu.co

G. Sánchez-Torres is with Systems and Computing Research Group, Faculty of Engineering, Universidad del Magdalena, Santa Marta, Colombia e-mail: gsanchez@unimagdalena.edu.co

with expressions that indicate some sentiment within the text [13][14][15]. On the other hand, there are more advanced approaches such as those based on supervised learning [16][17][18][19] and those that use models based on probabilistic inference [20][21][22].

Of all the above approaches, the vast majority do not considers the meaning of the words representing the aspects [23]. These are considered as simple ”tags” that are not located in the context of the opinion or in the domain of the entity to which it is referring. This article proposes an approach considering the meaning of the aspects and uses for it semantic techniques based on the ontologies, which have been successfully used in NLP tasks such as information extraction, disambiguation of the meaning of words, and automatic summarizing of texts, among others [24]. Additionally, the extraction of implicit aspects has not been addressed in depth in the Spanish language [25] and thus OM systems at the level of aspects lose coverage and effectiveness in the extraction process.

This article proposes a model that allows us to extract aspects (properties of an entity) from an opinion written in the Spanish language. The model is based on a domain ontology and semantic similarity techniques that contribute their semantic structures to the discovery of explicit aspects and double propagation techniques together with a co-occurrence matrix for the identification of implicit aspects.

III. Methodology

This article proposes a semantic model (SM) that makes it possible to extract explicit and implicit aspects from a text written in Spanish that represents a possible opinion about a specific entity. The model (see Figure 1) allows verification of whether a set of candidate aspects are written in the terminology of a specific domain with the help of a domain ontology [23] and a lexical database.

This semantic model is part of a complete OM system based on aspects known as AspectSA [25][26]. This system has a first preprocessing layer that takes a text in Spanish (opinion) and performs a process of lemmatization and labeling [27]. Therefore, the entry to the model is a set of tagged and lemmatized words represented as $C(WEL)$, which is analyzed by the semantic model that determines a set of aspects $C(A)$ identified as explicit and implicit.

A. Layer 1: Identification of candidate aspects

In the context of this proposal, a candidate aspect is a nominal expression (word or set of words) with a grammatical category ”noun” that is within the text of the opinion. That is, from an opinion, we identify a set of words (W), labeled (E) and lemmatized (L), and then determine the words W that have a grammatical category noun. For example, for the text ”Quien sea amante de la carne tiene una carta bastante amplia para elegir, aunque ayer no tenían chuletón”, the candidate aspects would be: ”amante”, “carne”, “carta” y “chuletón”, as shown in I.

B. Layer 2: Extraction of aspects with ontology

According to [28], the ontologies provide a structured and formal representation of knowledge, with the advantage of reusability and shareability, providing a common vocabulary defining a domain and the meaning of the concepts and the relationships between them. The above idea is used in this work to extract aspects of an opinion, taking advantage of the concepts, individuals (instances), and relationships of the ontology. Initially, the ontology is identified and selected depending on the language and domain being analyzed. We search for the candidate aspects in the ontology by comparing each with classes and individuals. Candidates that match the ontology are marked as explicit aspects.

For example, if we have an ontology that models the domain of hotels, such as Hontology (see Figure 2), this ontology contains concepts and individuals related to the domain [29]. If we have an opinion like “Mi estancia Hilton fue gratificante. Las habitaciones estuvieron estupendas”, the semantic model can initially identify that “habitación” is an aspect since it coincides with an ontology class. In

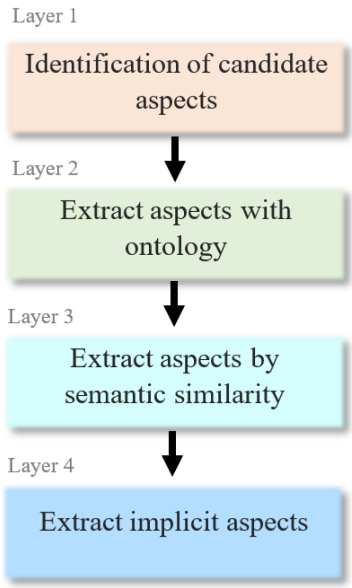

Fig. 1: Proposed semantic model.

TABLE I: Example of identification of candidate aspects.

Opinion	Tagged and Lemmatized	Candidate Aspects
Quien sea amante de la carne tiene una carta bastante amplia para elegir, aunque ayer no tenían chuletón	C(WEL) = (“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)	Amante, Carne, Carta, Chuletón

addition, the system can identify "Hilton" as it is probably an individual of the "Hotel" class belonging to the ontology.

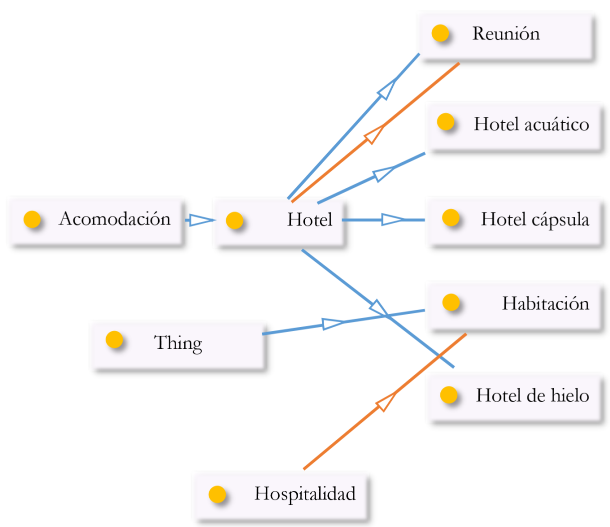

Fig. 2: An extract from the ontology *Hontology*.

The model allows the extraction of explicit aspects in other domains and other languages just by changing the ontology. This allows a system that implements the model to have high scalability. Algorithms 1 and 2 shown the procedure used for this process.

Algorithm 1: Getting classes and individual, getDataOntology

```
Data: ONTO ontology
Result: lists concepts and instances of ONTO
create COL collection
for all ONTO classes do
    class ← ONTO.next
    save class in COL
    get label from class in Spanish
    save label in COL
    instances ← class.getInstances
    for all instances do
        Individual ← instances.next
        save individual in COL
```

C. Layer 3: Extraction aspects by semantic similarity

After the previous process, the nouns representing the opinions that have not been found in the ontology are subjected to a process of finding the semantic similarity with the classes of the ontology (see Algorithm 3). In this proposal, the calculation of semantic similarity is based on the Wu and Palmer algorithm, which considers the position of concepts c_1 and c_2 in a taxonomy with respect to the position of the most specific common concept between the two $(c1, c2)$, see (Eq. 1)). It assumes that the similarity between two concepts is a function of the length and depth of the trajectory [30].

$$sim_{WP}(c_1, c_2) = \frac{2 * depth(Iso(c_1, c_2))}{len(c_1, c_2) + 2 * depth(Iso(c_1, c_2))} \quad (1)$$

Algorithm 2: Extraction of aspects, extractAspects

```
Data: list of concepts and instances COL, List of
      aspects OPINIONS
Result: lists of explicit aspects ASPECTS
create ASPECTS collection
for all opinion from OPINIONS do
    op ← OPINIONS.next
    for all words from op do
        wd ← OPINIOS.word
        if wd.label is NOUN then
            for all objects from COL do
                object ← COL.next
                if object.lemma == wd.lemma then
                    ASPECTS.add(wd)
```

To find the similarity, it is taken into account that the length $len()$ of the same concept is 0, $lso(c_1, c_2)$ is the common ancestor, and $depth(x)$ is the depth from the root taking into account that $depth(root) = 1$. For example, if we want to calculate the semantic similarity between two concepts such as "*Almuerzo*" and "*Cena*", based on Palmer distance and the taxonomy shown in Figure 3, then the depth from the root to the most common ancestor "*comida*" is equal to two (2), that is, $depth(lso(\text{"}Almuerzo\text{"}, \text{"}cena\text{"})) = 2$, the length is 2, that is, $len(\text{"}almuerzo\text{"}, \text{"}cena\text{"}) = 2$, and thus $sim_{wp}(\text{"}almuerzo\text{"}, \text{"}cena\text{"}) = 0.667$.

Algorithm 3: SemanticSimilarity

```
Data: list of opinions OPINIONS, list of explicit
      aspects ASPECTS, list of concepts and
      instances COL
Result: list of explicit aspects by similarity
        ASPECTSSIMIL
create ASPECTSSIMIL collection
for all opinions from OPINIONS do
    op ← OPINIONS.next
    for all words from op do
        wd ← op.next
        if wd.label in NOUN then
            if not(ASPECTS.contains(wd)) then
                for all objects from COL do
                    objects ← COL.next
                    simil ← similarityPalmer(wd,
                      object)
                    if simil > 0.65 then
                        ASPECTSSIMIL.add(wd)
```

Table II shows the semantic similarity calculated between the concept "*almuerzo*" (c_1) and the other concepts shown in the taxonomy of Figure 3.

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

TABLE II: Semantic similarity between two concepts c1 = *"almuerzo"*.

Concept	Len(c_1, c_2)	Depth(Iso(c_1, c_2))	Simwp(c_1, c_2)
Almuerzo	0	2	1
Comida	1	2	0.8
Cena	2	2	0.667
Alimento	2	2	0.667
Gastronomía	3	1	0.4
Verdura	4	2	0.5
Fruta	5	3	0.545
Manzana	6	4	0.571
Manzana Postre	7	5	0.588

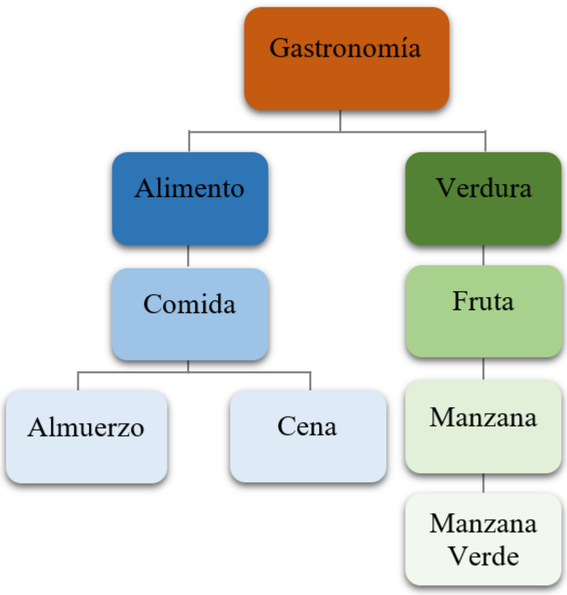

Fig. 3: An example of a gastronomy taxonomy.

D. Layer 4: Extraction of implicit aspects

In this article, for the extraction of implicit aspects in Spanish, the best characteristics obtained from the literature are taken in combination with the use of domain ontology. Double propagation techniques are used together with a co-occurrence matrix of explicit aspects and opinion words to determine possible implicit aspects [31][32][33].

The implicit ones are sought in those opinions or sentences where there is no explicit aspect. To assemble the co-occurrence matrix, the double propagation technique is used, starting with the candidate aspects based on the first level classes of the domain ontology.

The proposed process for the identification of implicit aspects entails: i) selection of a corpus of the defined domain, ii) definition of a seed list of opinion expressions, iii) definition of a seed list of possible explicit aspects from the ontology classes (first level), iv) the process of double propagation with seeds (opinion words and possible aspects) to find more opinion words and aspects that are affected by them, and v) the process of calculating the matrix of co-occurrence of aspects and expressions of opinion. Algorithms 4 and 5 show the procedure used for this process. It should be noted that the output of this component of implicit aspects is the explicit aspects related to the implicit aspects found in the opinion. For example, if the opinion is *"inmejorable"*, the component may give an explicit aspect related to *"comida"* for this opinion.

Algorithm 4: doblePropagation

```
Data: ontology ONTO, opinion corpus CORP
Result: list of aspects ASPECT, list of opinion
        expressions EXPR
out ← true
while out == true do
    create ASPECTS and EXPR collection
    out ← false
    for all classes from ONTO do
        class ← ONTO.next
        ASPECTS.add(class)
        for all opinion from CORP do
            opin ← CORP.next
            if opin.contains(class) then
                for all word from opin do
                    wd ← opin.next
                    if wd.label is VERB or
                     ADJECTIVE or ADVERB then
                        if !EXPR.contains(wd) then
                            EXPR.add(wd)
                            out ← true
    for all word from EXPR do
        wd ← EXPR.next
        for all opinion from CORP do
            opin ← CORP.next
            if opin.contains(wd) then
                for all word from opin do
                    wdop ← opin.next
                    if wdop.label is NOUN then
                        if !ASPECT.contains(wdop)
                         then
                            ASPECT.add(wdop)
                            out ← true
```

Algorithm 5: Obtaining co-ocurrence matrix

```
Data: list of aspects ASPECT, list of opinion
      expressions EXPR, opinion corpus CORP
Result: co-ocurrence matrix MATCO
for all aspects from ASPECT do
    asp ← ASPECT.next
    for all expressions from EXPR from op do
        exp ← EXPR.next
        for all opinion from CORP do
            opin ← CORP.next
            if opin.contains(asp) and
             opin.contains(expr) then
                cont ← cont+1
        MATCO[asp][exp] ← cont
```

IV. EXPERIMENTS AND RESULTS

To validate the proposed system, a series of experiments were carried out, taking as reference the corpus of task 5 referring to aspect-based OM (aspect-based sentiment analysis) of the 2016 edition of SemeEval (International Workshop on Semantic Evaluation), an organization that performs, by way of competition, continuous evaluations of computer systems of semantic analysis. Specifically, subtask 1 (SB1) was addressed in the Spanish restaurant domain [34].

Subtask SB1, in turn, is divided into three subtasks, called slots. Slot 2 consists of detecting the opinion target expression (OTE) of an E-A pair, that is, the linguistic expression used in the opinion to reference the entity (E) and the attribute (A). In this article, the OTE refers to the explicit or implicit aspect of the proposed model. For the experiments, the domain of restaurants in Spanish was taken. The training data with 2070 sentences and 627 texts and the evaluation data with 881 sentences and 268 texts were processed by the system. As an evaluation measure for slot 2, measure F1 was used, which was calculated from the precision and recall measurements.

The implementation of the model was done by building AspectSA software using Java technology, integrating different tools and libraries for the management of the Spanish language. To address the subtask of slot 2, the multilingual ontology ”Hontology” [29] was used, considering only the part in Spanish, and the most important characteristics of the ”Restaurant” ontology were adapted [35]. In addition, this ontology was extended by adding those instances that appeared in the training set provided for the subtasks.

For the calculation of semantic similarity, the multilingual lexical knowledge base MCR of wide coverage based on Wordnet was used. MCR integrates six different versions of the English Wordnet (from 1.6 to 3.0) and also Wordnets in Spanish, Catalan, and Italian, along with more than one million semantic relationships between concepts as well as semantic properties of different ontologies [36].

To find the implicit aspects, it is necessary to invoke two processes outside the AspectSA system. The first has as input a list of objects of ontology classes and uses the double propagation technique to find the nominal expressions whose labels are the adjectives, adverbs, and verbs that are related to the aspect. This double propagation technique is applied to an opinion corpus of more than $50,000$ opinions where the aspects to find the nominal expressions are considered first. When the process is over, we start with the nominal expressions found and look for a relationship with words that are substantive in the opinion. The process ends when there is no new look or new expression.

The second process takes the list of aspects and expressions and a co-occurrence matrix is constructed where the aspect appears along with the nominal expression in each corpus opinion.

The results of the task of extracting aspects of the AspectSA system in the domain of restaurants in the evaluation corpus for SemEval tasks are shown in Table III.

Table III shows the general results of the extraction of aspects: explicit with ontology ($column2$) and explicit with ontologies, similarity, and implicit ($column3$). In this table it can be seen that the recall is higher than the precision in both cases. This is an indication that the system for this domain correctly identifies many aspects and stops detecting only a few; however, the precision is lower because there are many false positives (many aspects were detected). In the same way, a high value is obtained for measure F1, mainly influenced by the recall.

TABLE III: Results for extraction of aspects in the evaluation corpus

Measure	Model without similarity	Full model
Precision	59.32	63.14
Recall	73.32	86.71
F1	65.58	73.07

The results obtained were compared with the final results of SemEval 2016 for the domain of restaurants, subtask SB1, and Spanish language (see Table IV). Column 2 of Table IV shows all the participants in the competition only in SB1, in the domain of restaurants (REST), and in the Spanish language (SP). The name of the device appears in the list followed by the letter U or C and then the measurement value. The letter C indicates that it is restricted only to the training data provided and the letter U indicates that it is not restricted but allows additional resources, such as lexical or training data. Table IV shows the measured values of $F1$ for the task evaluated. In the final part of the second column, the baseline is shown as the initial reference value.

TABLE IV: Results in the restaurant domain

Lang Dom Sub	Slot 2 - F1 SemEval	Slot 2 -F1 Model without similarity	Slot 2 - F1 Full model
SP REST SB1	GTI/C/68.515 GTI/U/68.387 IIT-T./U/64.338 TGB/C/55.764 basel./C/51.914	65.58	73.07

As can be seen, AspectSA (proposed system) obtains F1 values higher than the winners of the competition in the extraction of aspects. In the SemEval competition, the best results for slot 2 were obtained by the team GTI. Figure 4 shows the results of the slot2 of the Semeval competition and the AspectSA system. Here the AspectSA system considerably exceeds the best system in the competition by almost 5 points.

Analyzing the results of the extraction of aspects (slot 2), it should be noted that the use of domain ontology was vital for the identification of aspects, since this is an abstract model of a domain, where the concepts used are clearly defined and not simple dictionaries. By reusing a validated ontology in other tasks, it was possible to perform an extraction that took into account the meaning, allowing this knowledge to be exploited to improve the extraction performance.

Additionally, it should be noted that the semantic similarity method used in this work to address the extraction of aspects has contributed significantly to the improvement of the process. For the evaluation set, the F1 value of 65.58 has been increased, using only the ontology, to an F1 value of 73.07 using the ontology and semantic similarity.

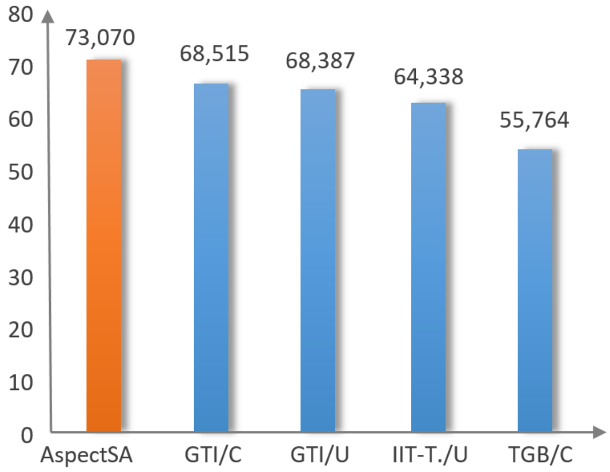

Fig. 4: Parts of the ontology used for the process of extracting aspects.

The presented system is highly scalable to any language and domain making small adjustments. For example, to work in the hotel domain, the system was adjusted with the original ontology ”Hontology” and the opinion corpus ”Hopinion” (http://clic..edu/corpus/hopinion) in Spanish, which contains around of $17,934$ reviews and 2,388,848 words, basically about hotels, from the TripAdvisor website. As there is no labeled evaluation corpus for this task, 120 different opinions were taken from the Web (Four experiments by Booking, TripAdvisor, Trivago and Expedia) in the hotels domain 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 by cross-validation 10 times. 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 V.

TABLE V: Results of experiments in the hotel domain

Web reviews	F1
TripAdvisor	89.18
Booking	85.72
Trivago	86.67
Expedia	87.01

You can see in Table V 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.

V. Conclusion

In this article, a semantic model was proposed for the extraction of explicit and implicit aspects in the Spanish language as a subtask of opinion mining. The system validation tests were carried out by means of a series of experiments, taking as reference the corpus of task 5 referring to aspects-based OM (aspect-based sentiment analysis) of the 2016 edition of SemEval. In the experiments, it was found that the proposed system obtained an F1 value of 73.07 in the process of extracting aspects, obtaining better results than the systems participating in SemEval.

Additionally, you will experiment in the hotel domain, where you will get a maximum F1 value of 89.18 for skin extraction using ten-fold cross-validation. The above, to prove the system easily scalable to other languages and domains, replacing the ontology and the corpus of opinions.

VI. Acknowledgments

This research was funded by the Universidad del Magdalena, Colombia.

References

[1] C. Henríquez and J. Guzmán, “Extracción de información desde la web para identificar acciones de un modelo de dominio en planificación automática,” *Ingeniare. Revista chilena de ingeniería*, vol. 23, pp. 439 – 448, 2015.

[2] F. A. Rocha Silva, “Analytical intelligence in processes: Data science for business,” *IEEE Latin America Transactions*, vol. 16, no. 8, pp. 2240–2247, 2018.

[3] J. Zhao, K. Liu, and L. Xu, “Sentiment analysis: Mining opinions, sentiments, and emotions,” *Computational Linguistics*, vol. 42, no. 3, pp. 595–598, 2016.

[4] B. Dalila, A. Mohamed, and H. Bendjanna, “A review of recent aspect extraction techniques for opinion mining systems,” in *2018 2nd International Conference on Natural Language and Speech Processing (ICNLSP)*, 2018, pp. 1–6.

[5] C. Henríquez, J. Guzmán, and D. Salcedo, “Minería de opiniones basado en la adaptación al español de anew sobre opiniones acerca de hoteles,” *Procesamiento del Lenguaje Natural*, vol. 56, pp. 25–32, 2016.

[6] C. Henríquez, F. Pla, l. F. Hurtado, and J. Guzmán, “Análisis de sentimientos a nivel de aspecto usando ontologías y aprendizaje automático,” *Procesamiento del Lenguaje Natural*, vol. 59, pp. 49–56, 2017.

[7] J. Steinberger, T. Brychcín, and M. Konkol, “Aspect-level sentiment analysis in Czech,” in *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Baltimore, Maryland: Association for Computational Linguistics, Jun. 2014, pp. 24–30. [Online]. Available: https://www.aclweb.org/anthology/W14-2605

[8] F. Xianghua, L. Guo, G. Yanyan, and W. Zhiqiang, “Multi-aspect sentiment analysis for chinese online social reviews based on topic modeling and hownet lexicon,” *Knowledge-Based Systems*, vol. 37, pp. 186 – 195, 2013.

[9] K. Schouten and F. Frasincar, “Finding implicit features in consumer reviews for sentiment analysis,” in *Web Engineering*, S. Casteleyn, G. Rossi, and M. Winckler, Eds. Cham: Springer International Publishing, 2014, pp. 130–144.

[10] K. Schouten and F. Frasincar, “Survey on aspect-level sentiment analysis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813–830, 2016.

[11] W. Zhang, H. Xu, and W. Wan, “Weakness finder: Find product weakness from chinese reviews by using aspects based sentiment analysis,” *Expert Systems with Applications*, vol. 39, no. 11, pp. 10 283 – 10 291, 2012.

[12] C. F. Samet, Y. Ezgi, Ö. Can, and E. Gulsen, “Tgb at semeval-2016 task 5: Multi-lingual constraint system for aspect based sentiment analysis,” in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics, Jun. 2016, pp. 337–341. [Online]. Available: https://www.aclweb.org/anthology/S16-1054

[13] G. P. Aitor, M. Cuadros, G. Rigau, and S. Gaines, “Unsupervised acquisition of domain aspect terms for aspect based opinion mining,” *Procesamiento del Lenguaje Natural*, 2014. [Online]. Available: https://www.redalyc.org/articulo.oa?id=515751574013

[14] S. Poria, E. Cambria, L.-W. Ku, C. Gui, and A. Gelbukh, “A rule-based approach to aspect extraction from product reviews,” in *SocialNLP*, 2014.

[15] A. Kumar, S. Kohail, A. Kumar, A. Ekbal, and C. Biemann, "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 the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics, Jun. 2016, pp. 1129–1135. [Online]. Available: https://www.aclweb.org/anthology/S16-1174
[16] D. Marcheggiani, O. Täckström, A. Esuli, and F. Sebastiani, "Hierarchical multi-label conditional random fields for aspect-oriented opinion mining," in *Advances in Information Retrieval*, M. de Rijke, T. Kenter, A. P. de Vries, C. Zhai, F. de Jong, K. Radinsky, and K. Hofmann, Eds. Cham: Springer International Publishing, 2014, pp. 273–285.
[17] M. Jabreel, F. Hassan, and A. Moreno, *Target-Dependent Sentiment Analysis of Tweets Using Bidirectional Gated Recurrent Neural Networks*. Cham: Springer International Publishing, 2018, pp. 39–55.
[18] A. G. Soujanya Poriaa, Erik Cambriab, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, pp. 42 – 49, 2016.
[19] H. Xu, B. Liu, L. Shu, and P. S. Yu, "Double embeddings and CNN-based sequence labeling for aspect extraction," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 592–598. [Online]. Available: https://www.aclweb.org/anthology/P18-2094
[20] S. Brody and N. Elhadad, "An unsupervised aspect-sentiment model for online reviews," in *Human Language Technologies*. Association for Computational Linguistics, 2010, pp. 804–812.
[21] L. Shu, H. Xu, and B. Liu, "Lifelong learning CRF for supervised aspect extraction," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. Vancouver, Canada: Association for Computational Linguistics, Jul. 2017, pp. 148–154.
[22] M. Shams and A. Baraani-Dastjerdi, "Enriched lda (elda): Combination of latent dirichlet allocation with word co-occurrence analysis for aspect extraction," *Expert Systems with Applications*, vol. 80, pp. 136 – 146, 2017.
[23] C. Henríquez and J. Guzmán, "Las ontologías para la detección automática de aspectos en el análisis de sentimientos," *Prospectiva*, vol. 14, pp. 90–98, 2016.
[24] C. Henríquez and J. Guzmán, "A Review of Sentiment Analysis in Spanish," *Tecciencia*, vol. 12, no. 22, pp. 35 – 48, 2017.
[25] C. Henríquez and E. Buelvas, "Aspectsa: Unsupervised system for aspect based sentiment analysis in spanish," *Rev Prospectiva*, vol. 17, no. 3, pp. 87 – 95, 2019.
[26] C. Henríquez, F. Briceño, and D. salcedo, "Unsupervised model for aspect-based sentiment analysis in spanish," *IAENG International Journal of Computer Science*, vol. 46, no. 3, pp. 430–438, 2019.
[27] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Computer Science*, vol. 17, pp. 26 – 32, 2013.
[28] I. Peñalver-Martinez, F. Garcia-Sanchez, R. Valencia-Garcia, M. Ángel Rodríguez-García, V. Moreno, A. Fraga, and J. L. Sánchez-Cervantes, "Feature-based opinion mining through ontologies," *Expert Systems with Applications*, vol. 41, no. 13, pp. 5995 – 6008, 2014.
[29] C. Marcirio, F. Larissa, and V. Renata, "Hontology: A multilingual ontology for the accommodation sector in the tourism industry," in *KEOD 2012 - Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, 2012, pp. 149–154.
[30] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*, ser. ACL '94. USA: Association for Computational Linguistics, 1994, pp. 133—138.
[31] M. Tubishat, N. Idris, and M. A. Abushariah, "Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges," *Information Processing Management*, vol. 54, no. 4, pp. 545 – 563, 2018.
[32] Y. Zhang and W. Zhu, "Extracting implicit features in online customer reviews for opinion mining," in *Proceedings of the 22nd International Conference on World Wide Web*, ser. WWW '13 Companion. New York, NY, USA: Association for Computing Machinery, 2013, pp. 103—-104.
[33] L. Sun, S. Li, J. Li, and J. Lv, "A novel context-based implicit feature extracting method," in *2014 International Conference on Data Science and Advanced Analytics (DSAA)*, 2014, pp. 420–424.
[34] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S. M. Jiménez-Zafra, and G. Eryiğit, "SemEval-2016 task 5: Aspect based sentiment analysis," in *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*. San Diego, California: Association for Computational Linguistics, Jun. 2016, pp. 19–30.
[35] J. Recio-Garcia. (2006) restaurant.owl. [Online]. Available: http://www.disi.unige.it/person/LocoroA/download/wilfontologies/restaurant.owl
[36] A. Gonzalez-Agirre and G. Rigau, "Construcción de una base de conocimiento léxico multilíngüe de amplia cobertura," *Linguamatica*, vol. 5, no. 1, pp. 13–28, 2013.