A novel deterministic approach for aspect-based opinion mining in tourism products reviews

Edison Marrese-Taylor [a], Juan D. Velásquez [a,*], Felipe Bravo-Marquez [b]

[a] Department of Industrial Engineering, Universidad de Chile, Av. República 701, P.O. Box: 8370439, Santiago, Chile
[b] Department of Computer Science, The University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

ARTICLE INFO

ABSTRACT

This work proposes an extension of Bing Liu's aspect-based opinion mining approach in order to apply it to the tourism domain. The extension concerns with the fact that users refer differently to different kinds of products when writing reviews on the Web. Since Liu's approach is focused on physical product reviews, it could not be directly applied to the tourism domain, which presents features that are not considered by the model. Through a detailed study of on-line tourism product reviews, we found these features and then model them in our extension, proposing the use of new and more complex NLP-based rules for the tasks of subjective and sentiment classification at the aspect-level. We also entail the task of opinion visualization and summarization and propose new methods to help users digest the vast availability of opinions in an easy manner. Our work also included the development of a generic architecture for an aspect-based opinion mining tool, which we then used to create a prototype and analyze opinions from TripAdvisor in the context of the tourism industry in Los Lagos, a Chilean administrative region also known as the Lake District. Results prove that our extension is able to perform better than Liu's model in the tourism domain, improving both Accuracy and Recall for the tasks of subjective and sentiment classification. Particularly, the approach is very effective in determining the sentiment orientation of opinions, achieving an *F*-measure of 92% for the task. However, on average, the algorithms were only capable of extracting 35% of the explicit *aspect expressions*, using a non-extended approach for this task. Finally, results also showed the effectiveness of our design when applied to solving the industry's specific issues in the Lake District, since almost 80% of the users that used our tool considered that our tool adds valuable information to their business.

1. Introduction

With the inception of the Web 2.0 and the explosive growth of social networks, enterprises and individuals are increasingly using the content in these media to make better decisions (Park & Kim, 2009; Zhu & Zhang, 2010). For instance, tourists check opinions and experiences published by other travelers on different Web platforms when planning their own vacations. On the other hand, for organizations, the vast amount of information available publicly on the Web could make polls, focus groups and some similar techniques an unnecessary requirement in market research.

However, due to the amount of available opinionated text, users are often overwhelmed with information when trying to analyze Web opinions. So far, many authors have tacked the problem of human limitation to process big amounts of information and extract consensus opinions (Velásquez & González, 2010) from a large number of sources relying on data-mining-based tools. Considering a similar problem, this work is an effort to create a tool that offers a set of summarization methods and help users digest in an easy manner the vast availability of opinions in the tourism domain. The core of our system is a novel extension of Bing Liu's aspect-based opinion mining methodology, which was developed by us in order to apply Liu's ideas to the tourism domain.

This extension is concerned with the fact that users refer differently to different kinds of products when writing reviews on the Web. Concretely, consider a *generic product*, which refers to the conceptual commodity produced by an industry (Smith, 1994). Most of the authors, including Kotler (2001), tend to classify these generic products using two categories, physical

* Corresponding author. Tel.: +56 2 2978 4834.
E-mail addresses: emarrese@wi.dii.uchile.cl (E. Marrese-Taylor), jvelasqu@dii.uchile.cl (J.D. Velásquez), fjb11@students.waikato.ac.nz (F. Bravo-Marquez).
URLs: http://wi.dii.uchile.cl/ (J.D. Velásquez), http://www.cs.waikato.ac.nz/~fjb11/ (F. Bravo-Marquez).

goods and intangible services. To the best of our knowledge, most of the existing works in this topic, including Liu's, are focused only on physical product reviews. In these kinds of reviews, users generally go straight to the point and talk directly about the product features they liked or did not like. Furthermore, few people will care about issues like who has designed or manufactured the product. However, for other kinds of products, different phenomena occur.

Works like (Cruz, Troyano, Enríquez, Ortega, & Vallejo, 2013) have already discussed the importance of the domain in the field of opinion mining. For instance, (Zhuang, Jing, & Zhu, 2006) indicates that when a person writes a movie review, he probably comments not only movie elements, but also movie-related people. However, few authors have focused into the field of tourism products like restaurants, which provide a physical good (the food) but also services in the form of ambience and the setting. A detailed study of on-line tourism product reviews revealed the most prominent features appearing on this domain, which we then capture and model in our extension. In general terms, we realized that users tend to *tell stories* about their experiences when writing these reviews, using longer and more complex sentences. The following example, taken from a real review in TripAdvisor, is intended to introduce the features that we will later focus on.

> "We had a lot of trouble finding the place, but after a while we finally made it. When we arrived to the hotel, it looked really good and only after trying several rooms we discovered the whole hotel was really mouldy in the interior. I barely had enough room to move around the 2 very small/short twin beds and the bathroom was smaller than most standard closets."

In the first place, a lot of sentences include multiple mentions of the product that is being reviewed or also of any of its features and components. On the other hand, a lot of sentences contain no opinions, also mentioning objects that do not correspond to attributes or components of the reviewed product. These sentences are usually explanations of the writer's experience and help to elaborate the *story* is being told. Finally, we realized that tourists might use many different and complex expressions to refer to the features or subcomponents of the reviewed product.

Therefore, the contributions of this paper are mainly three. First, to the best of our knowledge existing approaches do not address the special issues detected in the tourism domain, so we developed a model for aspect-based opinion mining that specially considers these features. This extension also included the development of new summarization and visualization methods that give insights about the customer preferences of each reviewed product. Our idea is based on the well known proposals of Lancaster in Lancaster (1966), which state that customer preferences about a product are intrinsically related to its features. The proposal is that discovering what these features are and defining how customers feel about these features will undoubtedly lead to a better comprehension of preferences, conceived as an evaluative judgment in the sense of liking or disliking an object (Scherer, 2005).

Secondly, as a result of the analysis of the domain, we created special corpora or datasets that help portraying the features of the mentioned domain. We also use these datasets for the evaluation of the proposed models for opinion aspect-based mining. Finally, our work also included the development of a generic architecture for an aspect-based opinion mining tool, which we used to create a prototype to analyze opinions from TripAdvisor in the context of the tourism industry in Los Lagos, a Chilean administrative region also known as the Lake District. Our system was intended to help users understand the attitude and the overall appreciation of Web users in the tourism domain by easily finding and extracting relevant subjective information from customer reviews published in TripAdvisor.

The rest of this paper is structured in the following manner. In first place, we discuss related state-of-the-art techniques and applications in Section 2. Later, in Section 3, we do a complete revision of Bing Liu's ideas, which served as inspiration of this work. Then, we introduce our extension in Section 4 and our system architecture in Section 5. After, we present the results of our experiments and application, in Section 6. Finally, Section 7 details conclusions and proposed future work.

2. Related work

Opinion mining or sentiment analysis comprises an area of NLP, computational linguistics and text mining, and refers to a set of techniques that deals with data about opinions and tries to obtain valuable information from them. As stated in Liu (2007), the literature offers two main approaches, aspect-based and non-aspect-based opinion mining. Aspect-based opinion mining techniques divide input texts into *aspects*, also called features or subtopics in literature, that usually correspond to arbitrary topics considered important or representative of the text that is being analyzed. The aspect-based approach is very popular and many authors have developed their own perspectives and models. Examples of them are (Archak, Ghose, & Ipeirotis, 2007; Decker & Trusov, 2010; Ku, Liang, & Chen, 2006; Lu, Zhai, & Sundaresan, 2009; Popescu & Etzioni, 2005; Titov & McDonald, 2008; Zhao & Li, 2009; Zhuang et al., 2006).

Based on an extensive revision of the state-of-the-art approaches and tools, we concluded that Bing Liu's ideas were probably the most comprehensive models on the topic of aspect-based opinion mining. For that reason, his ideas were used here by us as inspiration. In general, our work is based on the ideas summarized by Liu in Liu (2007), which includes a review of the state-of-the-art models, with special attention to his ideas. Most these ideas had already been discussed in the corresponding papers by Liu and his colleagues. Our approach is different from Liu's ideas since it is domain focused; intended to perform well with tourism product reviews. Other reviews of the state-of-the-art opinion mining techniques can be found in Kim, Ganesan, Sondhi, and Zhai (2011), Pang and Lee (2008) and Marrese-Taylor, Rodríguez, Velásquez, Ghosh, and Banerjee (2013).

Other related work includes (Xu, Cheng, Tan, Liu, & Shen, 2013), which proposes an approach for aspect-based opinion mining based on modified versions of Latent Dirichlet Allocation (LDA), similar to what is proposed in the pioneer paper (Titov & McDonald, 2008). These approaches are unsupervised topic-based document modeling techniques, which model an input document as a mixture of topics. A good example of this proposal can be found in Dueñas-Fenández, Velásquez, and LHuillier (2014), where authors present a framework for trend modeling and detection on the Web, based on the fusion of freely available information. In this context, our work lies on a radically different paradigm, as the former consists in identifying the aspects reviewed in a piece of text based on a bag-of-words model of the document, rather than extracting individual feature mentions and their related opinions (Cruz et al., 2013). Therefore, our work is not directly comparable to these kind of works.

On the other hand, it's also possible to mention (Cruz et al., 2013), which analyzes the importance of the domain in opinion mining. On the paper, the authors show that different topics have completely different features and issues. They also developed a system that by the means of human intervention by generating annotated corpora for each domain, is capable of performing well

across different domains. Regarding this, our work acknowledges the differences between domains that is discussed in the paper, but rather than proposing a general model that works for all the domains, we focus on the tourism domain in order to solve its specific issues. Also, our system does not require any training datasets and only a small amount of human support.

Finally, one last related topic is the set of so-called concept-level sentiment analysis approaches. These approaches focus on a semantic analysis of text through the use of Web ontologies or semantic networks, which allow the aggregation of conceptual and affective information associated with natural language opinions (Cambria, 2013; Havasi, Cambria, Schuller, Liu, & Wang, 2013a, Havasi, Cambria, Schuller, Liu, & Wang, 2013b). The concept-level approach is directly related to the bag-of-concepts model which is usually considered to he able to represent semantics associated with natural language text much better than bags-of-words. Current approaches of this kind mainly leverage on existing affective knowledge that helps understanding the semantics behind bag-of-concepts, such as WordNet. Indeed, (Cambria, Poria, Gelbukh, & Kwok, 2014) presents an API for concept-level sentiment analysis which provides semantics and sentics associated with 15,000 natural language concepts. Concept-level also includes high-level tasks such as domain adaptation, opinion summarization and multimodal sentiment analysis – analysis based on linguistic, audio, and visual features. In this context, our work is based only on linguistic features and does not use any external source of knowledge.

Regarding existing applications of opinion mining, since the topic has attracted the attention of many research fields, many tools exist so far. A considerable number of these applications consider Twitter as a source of opinionated documents, such as *Sentiment 140*[1] and *TweetFeel*.[2] On the other hand, *Socialmention*[3] offers a social media search and analysis platform that aggregates user-generated content from different social media sources. Our approach is different from all these applications since it is aspect-based and analyzes opinions at the sentence level.

In addition, there are a significant number of applications that mine sources that contain product reviews, such as the mentioned TripAdvisor and VirtualTourist (for tourism products) or Amazon and C|Net. Examples of these applications are the Lexalytics Salience Engine[4] and Nebular (Palakvangsa-Na-Ayudhya, Sriarunrungreung, Thongprasan, & Porcharoen, 2011). These applications process opinionated documents and generally offer text summaries as output, lacking other visualization methods. In this area, our tool is different since it offers novel and intuitive graphic summaries of opinions. These summaries are intended to provide users a way of processing the vast amount of information available in social media about tourism products. Lastly, we also found OpinionEQ,[5] which offers an approach that seems very similar to ours. However, OpinionEQ is not proposed as an application but rather as a service.

3. Background

In this section, we explain Bing Liu's models and ideas, which served as inspiration for our approach. The intention is to make it easier to understand how our proposals extend Liu's models and how they are tailored for the tourism domain. To start, let us first discuss Liu's approach in general terms. We see that Liu proposes that opinions are 5-tuples (Liu, 2007), composed of the following parts.

(1) An *entity*: Proposed to denote the opinion objective or, in other words, what is being evaluated by the opinion. An *entity* can contain a set of components and attributes and, similarly, each *entity* component can have its own subcomponents and attributes. Finally, an *entity* can be decomposed into a tree or hierarchy of subattributes and subcomponents.

(2) An *aspect*: Because it is difficult to study an *entity* at an arbitrary hierarchy level, this hierarchy is simplified to one or two levels, denoting as *aspect* every component or attribute of the *entity*. In this way, the root of the hierarchy or tree becomes the *entity* itself, each leaf is an *aspect* and links are part-of relationships.

(3) The *Sentiment orientation*, considering that opinions express a *positive* or *negative* sentiment about what they evaluate.

(4) The *Opinion holder*, which corresponds to the user (a person, an enterprise, etc.) that gives the opinion.

(5) *Time*: Time and date when the opinion was given.

In this manner, opinions are considered to be a positive or negative view, attitude, emotion or appraisal about an *entity* or an *aspect* of that *entity* from an *opinion holder* in a specific time. The following concepts are also introduced:

- *Entity expression*: Corresponds to the actual word or phrase written by the user to denote or indicate an *entity*. As a result, *entities* are then generalizations of every *entity expression* used in the analyzed documents, or a particular realization of an *entity expression*. In Liu (2007) this concept is called *entity name*.
- *Aspect expression*: As for an *entity expression*, the *aspect expression* is the actual word or phrase written by the user to denote or indicate an *aspect*. Thus, aspects are also general concepts that comprise every *aspect expression*. They are called *aspect names* by Bing Liu.

It is then possible to define a model of an *entity* and a model of an opinionated document. An *entity* e_i is represented by itself as a whole and a finite set of *aspects*, $A_i = \{a_{i1}, a_{i2}, \ldots, a_{in}\}$. The *entity* can be expressed with any one of a finite set of *entity expressions* $EE_i = \{ee_{i1}, ee_{i2}, \ldots, ee_{is}\}$. Each *aspect* a_{ij} of A_i of *entity* e_i can be expressed by any one of a finite set of *aspect expressions* $AE_{ij} = \{ae_{ij1}, ae_{ij2}, \ldots, ae_{ijm}\}$. On the other hand, an opinionated document $d_k \in D$ contains opinions on a set of *entities* $e_1, e_2, \ldots, e_r$ from a set of opinion holders $h_1, h_2, \ldots, h_p$. The opinions on each *entity* e_i are expressed on the *entity* itself and a subset A_{ik} of its *aspects*.

Kim et al. (2011) gives a good review of historical and state-of-the-art aspect-based developments. The authors indicate that the process is commonly made up of three distinct steps, which are also considered by Liu.

3.1. Aspect identification

This stage aims to find and extract important topics in the text that will then be used to summarize. In Hu and Liu (2004b), Hu and Liu present a technique based in NLP and statistics. In their proposal, part-of-speech (POS) tagging and syntax tree parsing (or chunking) are used to find nouns and noun phrases or NPs. Then, using frequent itemset mining, the most frequent nouns and NPs are extracted. The extracted sets of nouns and NPs are then filtered using special linguistic rules. These rules ensure that the terms inside those aspects that are composed of more than one word are likely to represent real objects together and also eliminate redundant aspects. They also extract non-frequent aspects using an approach by finding nouns or NPs that appear near to opinion words with high frequency. This approach does not extract adjectives or any other kind of non-object aspects.

3.2. Sentiment prediction

The next phase is sentiment prediction, to determine the sentiment orientation on each *aspect*. Ding, Liu and Yu offer a lexicon and rule-based approach in Ding, Liu, and Yu (2008). This method relies on a sentiment word dictionary that contains a list of positive and negative words (called *opinion words*) that are used to match terms in the opinionated text. Also, since other special words might also change the orientation, special linguistic rules are proposed. Among others, these rules consider negations words *"no"* or *"not"* and also some common negation patterns. However, despite how simple these rules might appear, it is important to handle them with care, because not all occurrences of such rules or word apparitions will always have the same meaning. In this context, rules developed by Ding, Liu and Yu include an aggregation score function to determine the orientation of an aspect in a sentence combining multiple *opinion words*. This function will be explained in detail in Section 4, since it will be used and extended by us.

3.3. Summary generation

The last step is summary generation, to present processed results in a simple manner. In this context, defined opinion quintuples are a good source of information for generating quantitative summaries. In particular, Liu defines a kind of summary called aspect-based opinion summary (Hu & Liu, 2004a, 2006), that consists of bar charts that show the number of positive and negative opinions about every *aspect* of one *entity*. Liu, Hu, and Cheng (2005), also proposes that the bar charts could be used to compare a set of selected products, showing the set of all aspects of the chosen products in the chart. In this case, each bar above or below the *x*-axis can be displayed in two scales: (1) the actual number of positive or negative opinions normalized with the maximal number of opinions on any feature of any product and (2) the percent of positive or negative opinions, showing the comparison in terms of percentages of positive and negative reviews.

4. Proposed extension

Our extension, based on the work of Marrese-Taylor, Velásquez, Bravo-Marquez, and Matsuo (2013), takes Liu's methods as a basis and considers the same set of structured steps mentioned in Section 3. Here, we discuss issues on each one of the tree steps and explain our own approach in the context of tourism product reviews.

4.1. Aspect expression extraction

As defined by Liu, *aspects* do not directly appear in a text but they exist in the manner of *aspect expressions*. Accordingly, when trying to apply Liu's opinion model to extract opinions from real data, concepts can be somewhat confusing or unclear. It is also unclear how *aspects* that appear more than once in a document are managed. Having noticed these issues, a model to build opinion tuples from an opinionated document has been developed here.

To make things simpler, consider a set of opinionated documents $D_i = \{d_{i1}, d_{i2}, \ldots, d_{im}\}$ about only one *entity*, e_i. This seems a realistic assumption since opinions are usually available in the form of product reviews on the Web. Then, each opinionated document will correspond to a review or opinion given by holder h_k in time t_k. Let S_{ik} be the set of all sentences in d_{ik}, with $S_{ik} = \{s_{ij1}, s_{ij2} \ldots, s_{ijn}\}$. Opinions on e_i in d_{ik} will be expressed on the *entity* itself and on a subset A_{ik} of its *aspects*. Similarly, each *aspect* of A_{ik} will appear on d_{ik} as a set of *aspect expressions* AE_{ijk},

subset of AE_{ij}. The entity e_i will appear as a subset of different *entity expressions* $EE_{ik} \subseteq EE_i$. Thus, the set EX_{D_i} is defined as the set of all *aspect expressions* of all *aspects* and all *entity expressions* appearing in D_i. A sentence is related to one *aspect expression* or *entity expression* only if it appears in that sentence. Next, sentiment orientation needs to be determined for each pair (ex, s) only if any *aspect expression* or *entity expression* appears on it. After determining sentiment orientation, h_k and t_k of the corresponding document d_{ik} should simply be added in order to build each opinion tuple.

On the other hand, Liu's proposal indicates that it seems reasonable that frequently used nouns in product reviews are usually genuine and important *aspects expressions* because when people comment on different *aspects* of a product, the vocabulary that they use usually converges. Nevertheless, two main reasons explain the fact that many different expressions could indicate the same concept, particularly in the tourism domain:

- The economy principle in languages (Vicentini, 2003) indicates that they try to say a lot using few words. For example, the sentence *"The hotel has good wifi."* corresponds to a lexicalization, where the original expression, *"The hotel has good Internet access through wifi."*, is shortened according to the economy principle.
- Each language presents systems that organize its concepts, also pursuing simplification. For that reason, many words in English (as in all other languages) simply are hyponyms of a determined hypernym. A hyponym is a word or phrase whose semantic field is included within that of another word, its hypernym. For instances, scarlet, vermilion, carmine, and crimson are all hyponyms of red (their hypernym), which is, in turn, a hyponym of color (Fromkin, Rodman, & Hyams, 2010).

In practice, finding the *aspects* that are evaluated in a set of opinionated documents is a really complex task. In fact, detecting *aspect expressions* from a set of documents with opinions should be a completely different task than defining or finding the real *aspects* in them, because the amount of possible expressions appearing in a text is really huge. Regarding this, we have already said that in tourism product reviews several expressions are in fact used.

Another issue found in Liu's proposals is related to the concepts of sentence and *word distance*, that although widely used, are not clearly defined. Despite deeper linguistic analysis, here we will define a sentence as an ordered set of tokens, including words and punctuation. One token that appears in two different positions must be considered twice, as the positions where they appear are distinct. In other words, a sentence S will correspond to a set of unique tuples (*token, position*). Positions can only be in $\mathbb{N} \cup \{0\}$ and the difference between two adjacent components must be 1. As such, the concept of *word distance* between two elements of sentence S will correspond to the difference of the positions of the two tokens in S.

$$WD(t_a, t_b) = |position(t_a) - position(t_b)|t_a, t_b \in S \qquad (1)$$

As $WD(t_a, t_b)$ (Word Distance) is simply the absolute value of the difference between numbers in $\mathbb{N} \cup \{0\}$, *Word Distance* (t_a, t_b) is a metric on the set S as it satisfies the conditions of non-negativity, identity of indiscernibles, symmetry and triangle inequality. Note that the minimal distance between 2 elements in S is 1, and it occurs between adjacent elements. The maximum distance corresponds to $|S| + 1$.

Despite these definitions and formalizations, in this work we focused on the task of determining the sentiment orientation at the aspect level, so here we merely apply the technique developed by Hu and Liu in Hu and Liu (2004b) to extract frequent aspects. In other words, in an effort to make the rest of the analysis simpler,

we will be considering *aspects expressions* to be only nouns or sets of nouns which we call *explicit aspects expressions*. We will not extract *implicit* nor not-frequent *aspect expressions*.

4.2. Determination of the opinion orientation

Taking the work of Ding et al. (2008) as inspiration, a set of rules to determine the sentence orientation was developed, always considering *opinion words* as a basis.

4.2.1. Word orientation rules

In first place, we need to determine the orientation of each word in a sentence. In order to do so, we propose Algorithm 1. The algorithm applies a set of linguistic rules, which are explained below.

Algorithm 1. Word orientation

1: **if** word is in opinion_words **then**
2:   mark(word)
3:   *orientation* ← Apply Opinion Word Rule(marked_word)
4: **else**
5:   **if** word is in neutral_words **then**
6:     mark(word)
7:     *orientation* ← 0
8:   **end if**
9: **end if**
10: **if** word is near a too_word **then**
11:   *orientation* ← Apply Too Rules(orientation)
12: **end if**
13: **if** word is near a negation_word **then**
14:   *orientation* ← Apply Negation Rules(orientation)
15: **end if**
16: **return** orientation

- Word rules: Positive *opinion words* will intrinsically have a *score* of 1, denoting a normalized positive orientation, while negative ones will have associated a *score* of −1. Every noun and adjective in each sentence that is not an *opinion word* will have an intrinsic *score* of 0 and will be called *neutral word*.
- Negation rules: A negation word or phrase usually reverses the opinion expressed in a sentence. Consequently, *opinion words* or *neutral words* that are affected by negations need to be specially treated. Three rules must be applied: Negation Negative →Positive, Negation Positive → Negative and Negation Neutral → Negative. Negation words and phrases include: *"no"*, *"not"*, *"never"*, *"n't"*, *"dont"*, *"cant"*, *"didnt"*, *"wouldnt"*, *"havent"*, *"shouldnt"* (misspellings are here intentional). Also, some negation patterns are considered, including *stop + vb-ing*, *quit + vb-ing* and *cease + to + vb*.
- Too rules: Sentences where words *"too"*, *"excessively"* or *"overly"* appear, are also handled specially. When an *opinion word* or a *neutral word* appears near one of the mentioned terms, denoted *too words*, its orientation will always be Negative (*score* = −1).

4.2.2. Aspect orientation rules

Having mentioned rules that help in determining each word orientation in a sentence, it is now explained how all these orientations should be combined to determine the final orientation of a sentence on a particular *aspect*. Our proposal is summarized in Algorithm 2 and it only considers words marked as *opinion words* or *neutral words*, which we call *marked words*, as they are the only ones that will provide the orientation for each sentence. The detailed process is explained below.

Algorithm 2. Opinion orientation

1: **if** but_word is in sentence **then**
2:   orientation ← Opinion Orientation(aspect,marked_words,but_clause)
3:   **if** orientation ≠ 0 **then**
4:     **return** orientation
5:   **else**
6:     orientation ← Opinion Orientation(aspect,marked_words,not but_clause)
7:     **if** orientation ≠ 0 **then**
8:       **return** −1 × orientation
9:     **else**
10:       **return** 0
11:     **end if**
12:   **end if**
13: **else**
14:   **for all** aspect_position in aspect **do**
15:     **for all** aspect_word in aspect_position **do**
16:       **for all** word in marked_words **do**
17:         suborientation += $\frac{Word\_Orientation(word)}{WD(aspect\_word,word)}$
18:       **end for**
19:       orientation += suborientation
20:     **end for**
21:     final_orientation += orientation
22:   **end for**
23:   **if** final_orientation > 0 **then**
24:     **return** 1
25:   **else**
26:     **if** final_orientation < 0 **then**
27:       **return** −1
28:     **else**
29:       **return** 0
30:     **end if**
31:   **end if**
32: **end if**

- Aspect words aggregation rule: Let *s* be a sentence that contains the set of *aspect expressions* $A = \{a_1, \ldots, a_m\}$, each one of them appearing only one time in *s*. Also, let AW_i be the set of words that comprise aspect a_i, where $AW_i = \{aw_{i1}, aw_{i2}, \ldots aw_{in}\}$. Each aw_{ij} will be called *aspect word* and it will correspond to an *aspect expression* a_i. If *scores* for each *opinion word* and *neutral word* in *s* are known, *score* for each aw_{ij} in *s* is given by the following aggregation function:

$$score(aw_{ij}, s) = \sum_{ow_j \in s} \frac{score(ow_j)}{WD(ow_j, aw_{ij})} \qquad (2)$$

where ow_j is an *opinion word* or *neutral word* in s, $WD(ow_j, aw_{ij})$ is the word distance between the aspect word aw_{ij} and the opinion word ow_j in *s*. Line 17 implements this formula in Algorithm 2. We take this function from Ding et al. (2008); however, their proposition lacked an explanation of how the function should be applied to *aspect expressions* that are composed of more than one word (which we call *compound*). We have seen that in tourism product reviews some *aspect expressions* are in fact *compound*. For instance, in the sentence *"The hotel had a poor view of the beautiful lake."* an *aspect expression* that should be extracted by Liu's algorithms is *lake view*. However, Liu's proposal does not explain how the orientation on this *aspect* should be obtained in the sentence. In order to consider these cases, we propose that the formula should not be used for each *aspect expression*

but rather for each word in each expression. These orientations are aggregated according to the next rule.

- Aspect aggregation rule: For each *compound aspect expression* a_i in *s*, its orientation will be calculated considering the *scores* of all the words that compose it, $aw_{ij} \in AW_i$, according to the following equation, which is implemented in line 19 of Algorithm 2.

$$score(a_i, s) = \sum_{aw_{ij} \in AW_i} score(aw_{ij}, s) \quad (3)$$

- Position aggregation rule: We have also seen that in tourism product reviews *aspect expressions* could appear more than once in a sentence. This case is not covered by Liu's proposals, but here we need a method to cover these cases. Supposing that a_i appears *t* times in *s* and knowing the score of each *aspect expression* appearance $a_i^k, k \in \{1, 2, \ldots, t\}$, we propose that the final score of a_i, or $fscore(a_i, s)$, should be calculated by simply adding the values of the scores of all the a_i appearances in *s*, according to the following equation.

$$fscore(a_i, s) = \sum_{k=1}^{t} score(a_i^k, s) \quad (4)$$

The formula appears in line 21 of Algorithm 2. Note that when a_i only appears one time in $s, fscore(a_i, s) = score(a_i, s)$. Finally, lines 23–31 show how the orientation is computed according to the *fscore* of each *aspect expression*. If $fscore(a_i, s)$ is positive, the opinion is considered positive on a_i (lines 23 and 24) and if it is negative, the opinion is considered negative on a_i (lines 26 and 27). If none of these cases occur, the sentence is considered neutral (line 29).

- But clauses rules: We use exactly the same rule proposed in Ding et al. (2008). This rule states that when a *but word b* (including the word *but* or any synonym) appears in sentence *s, s* must be broken into two segments, the one before and the one after *b*. If the orientation of any *aspect word* aw_{ij} appearing in the sentence segment after *b* is zero, its orientation should then be determined using the segment before *b*, but assigning the opposite result. We realized that a little ambiguity existed since in some of these cases aw_{ij} could appear outside of the considered segment. Here, we simply propose that aw_{ij} must be added at the final position of the corresponding segment in order to avoid the consistency issue. Lines 1–12 in Algorithm 2 apply this rule.

4.3. Summarization

Liu's proposal seems fairly simple and effective for summarizing opinions. However, it lacks a robust way of measuring the importance of each evaluated *aspect*. In Hu and Liu (2004a), *aspects* are ranked according to the frequency of their appearances in the reviews, but it is also declared that other types of rankings are also possible, like ranking *aspects* according to the number of reviews that express positive or negative opinions. Here, we attempt to measure the importance of each *aspect* simultaneously using the amount of positive and negative opinions of it. We also use that measure to rank *aspects*. The underlying assumption is that an *aspect* that has a lot of positive and negative opinions will be more important, since the high number of opinions of both orientations might indicate that customers are very interested in that *aspect*. In this way the total number of times that an *aspect* appears is not only considered in measuring importance, but also the dispersion in the number of positive and negative opinions. Let P_i and N_i be the number of positive and negative opinions on aspect $a_i, i \in \{1, \ldots, n\}$. Then, $PScore_i$ and $NScore_i$ will be the min-max

normalized values of P_i and N_i, respectively. With this, we calculate the standard deviation of these *scores* using:

$$AVScore_i = \frac{PScore_i + NScore_i}{2} \quad (5)$$

$$STDScore_i = \sqrt{\frac{(PScore_i - AvScore_i)^2 + (NScore_i - AvScore_i)^2)}{2}} \quad (6)$$

We define our new measure for each *aspect expression* a_i, called *Relative Importance*, as the min-max normalized value of its $STDScore_i$. We propose that aspect-based summaries should include bar charts and a table that shows the actual values of $PScore_i, NScore_i$ and *Relative Importance* for each *aspect expression*. As we will see in Section 6, since we plan to show the opinion mining results to users in our system, this measure will play an important role in assuring that the information we show is clean and appears with little noise.

5. System architecture

As introduced in Marrese-Taylor, Velásquez, and Bravo-Marquez (2013), our system was designed using a modular programming paradigm. Fig. 1 shows the proposed architecture. The main functionalities are described in the following paragraphs.

The Data Collection Module (DCM) is in charge of obtaining opinions from a set of given Web sources. This module simply consists of a set of Web crawlers, which must be source-specific. The crawlers parse HTML webpages containing opinions and pre-process the results, generating comma-separated CSV files containing the downloaded opinionated documents.

The Opinion Mining Module (OMM) implements the proposed aspect-based opinion mining algorithms on a given set of opinionated documents. Each opinionated document is separated into sentences, which are then split into tokens; POS tagging and syntactic chunking methods are then applied. Two different tasks need to be performed, *aspect* extraction and orientation determination, for which two sub-modules are included:

- Aspect extraction sub-module: in charge of applying the *aspect* extraction algorithm to a set of POS-tagged sentences. As we already said, this algorithm is based on (Hu & Liu, 2004b), which uses the most frequent nouns and NPs to extract aspects.
- Orientation determination sub-module: This sub-module applies the algorithms presented in Section 4 to determine the orientation of an opinion on a given aspect. It also extracts the set of adjectives that appeared near each aspect.

The Results Visualization Module (RVM) is the visible portion of the application and interacts directly with the user. Users can give opinion data to the system which can then be used to apply the opinion mining process. Results include the following features:

- Aspect-based summaries: Bar charts, in which each bar measures the number of positive and negative mentions of each attribute or component of one product. Bars are initially sorted according to *Relative Importance*.
- Adjective bubble charts: Nearby adjectives in all sentences where an *aspect* appears are shown in a bubble chart. The size of each bubble counts the times that each adjective is used to describe the aspect.
- Original opinions: A list of all original sentences is also displayed in an ad-hoc manner, separating them into positive or negative.

The system also provides a tagging interface that helps users to extract opinions from the opinionated documents and alter the algorithm's results. This functionality appears in a special menu
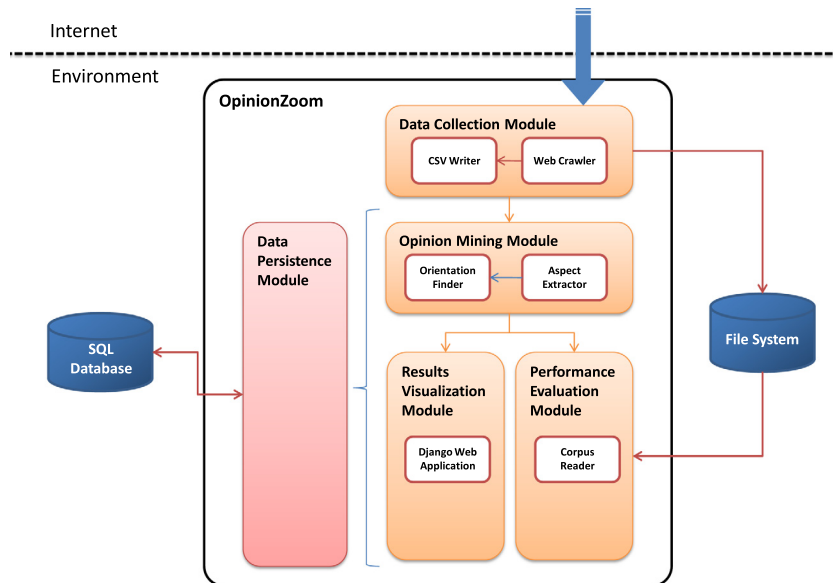
Fig. 1. General design of our system.

```
line
  1   [c1][s1] place[+], comfort[+][u], location[+][u] ### good place to stay at
      the end of a long flight in that it is very comfortable, with many facilities,
      and in the town , by the shore .
  2   [c1][s2] ### however, puerto montt is not the best of places to explore.
  3   [c1][s3] ### a better place is puerto varas which is just as near to the
      airport and has far more attractions.
  4   [c1][s4] ### could not fault this aparthotel.
  5   [c2][s1] hotel[+] ### my fiance and i spent three nights here in march
      2012 and it's a sweet, quaint hotel.
  6   [c2][s2] reservation[-] ### with that said, i called in early february 2012
      to make a reservation and it got lost/misplaced.
```

Fig. 2. Example sentences of an annotated corpora.

that does not interfere with the rest of the specifications. In addition, after applying the opinion mining algorithms, the system offers an interface that lets users see the list of the extracted aspect and select the ones that he really wishes to save. We included these two functionalities to receive relevance feedback from our users. Thus, choices and operations performed by users are stored and then used to improve the system performance.

The Performance Evaluation Module (PEM) is in charge of delivering a set of indexes that evaluates the performance of the opinion mining algorithms. In order to do this, the system allows users to elaborate and then provide specially annotated corpora, following the structure that appears in Fig. 2. To facilitate the annotation process, guidelines and examples are also offered. The annotation technique follows the spirit of what Liu proposes in Ding et al. (2008) and Hu and Liu (2004a). As a result, three tasks can be evaluated by comparing the extraction process results with the provided corpora: (1) Explicit aspect extraction, to measure the effectiveness of the explicit aspect extraction algorithm, (2) Subjectivity classification, to evaluate the effectiveness of opinion sentence extraction and (3) Sentiment classification, to measure the accuracy of the orientation prediction of each *aspect expression* in each sentence (ex, s), for the positive class. We believe that the service provided by this module is crucial when trying to understand the usefulness of the system within a particular topic or domain. To the best of our knowledge, this represents an important difference between ours and other existing tools.

Finally, the Data Persistence Module or DPM manages all the database operations and constitutes a model layer for the whole system. The data layer is implemented using two relational models, which support all the data that needs to be stored.

6. Experiments and industry application

In this section, we show a real case application where the proposed design was implemented using Python. The application encompasses the situation in the Lake District, where tourism operators lack tools to understand what their customers want or need. We also used our application and the data we collected to generate datasets (linguistic corpora) to evaluate the performance of the opinion mining algorithms implemented in the OMM. Our study used the *NLTK*[6] libraries for the NLP tasks in the OMM and the *Django Framework*[7] for the RVM.

6.1. Algorithm performance evaluation

In the first place, using the DCM, we downloaded all the reviews from hotels and restaurants originally written in English about the *Lake District* in TripAdvisor. We obtained a total of 1435 reviews

---

[6] http://nltk.org.
[7] https://www.djangoproject.com.

Table 1
Corpora details.

Corpus	Hotels	Restaurants
Reviews	100	100
Total sentences	789	470
Opinion sentences	609	368
Opinion sentences/sentences	77.19%	78.3%

Table 2
Detail on *aspects* found in corpora.

Aspect type	Hotels corpus		Restaurants corpus	
	Number	Percentage	Number	Percentage
Explicit	229	73.87	161	67.93
Explicit and implicit	30	9.68	26	10.97
Implicit	51	16.45	50	21.1
Total	310	100	237	100

and saved them in two different CSV files, as defined in the design. In order to generate the annotated corpora to evaluate the performance of the algorithms, we randomly selected 100 restaurant and hotel reviews. Later, each review was tokenized into sentences using the unsupervised machine learning algorithm proposed in Kiss and Strunk (2006). Finally, each sentence was manually annotated following our guidelines (for details see our corpora material[8]). Sentences that seemed ambiguous or really difficult to tag were discussed with a second human annotator, an expert in linguistics. Once an agreement was achieved, the sentence was tagged according to that agreement. This marks an important difference between this study and other tagging procedures commonly carried out in literature, where different annotators tag the same corpus separately and only once the annotation procedure has finished are different results of the same corpus compared to define the final choice. This different approach was used here due to time constraints, since it seemed more efficient and was worth trying as a contribution to research in this field.

Table 1 gives a general description of the generated corpora. In both cases, almost 80% of the sentences contained opinions. This shows that opinionated sentences represent an important fraction of the total sentences, which somewhat validates the use of TripAdvisor as a source of opinions for tourism product reviews. Nevertheless, as expected, non-opinionated sentences are also a considerable number, consequently introducing noise into the opinion-extraction process.

Table 2 gives details about the *aspect expressions* that were manually extracted. Following our notation, we call those expressions that appear in the manner of nouns or NPs in a sentence explicit *aspect expressions* and implicit *aspect expressions* to all other cases. Results show that in both corpora explicit *aspect expressions* are the most common ones, representing around 70% of all the extracted expressions. When some *aspects expressions* appear in both an explicit and implicit manner, they were considered as explicit. On the other hand, extracted *aspect expressions* that are purely implicit are also an important number, being almost 20% in both cases. A simple review showed that most of these *aspects* were indicated by adjectives.

A further analysis of the datasets consisted in finding the best-fitting distribution of the number of sentences for each case. Fig. 3 shows the charts with the best fitting discrete distributions and their parameters, obtained using maximum-likelihood estimation (MLE). As can be seen, for both cases the best-fitting distribution was Negative Binomial. This shows that the tourism product

reviews on the Web present strong similarities regardless of the specific product that is studied. The fact that EMV parameters were slightly different for each case was consistent with the differences in the number of sentences for each case (see Table 1) that we had observed in our first analysis.

We implemented all the PEM specifications and then evaluated how the proposed opinion mining algorithms perform when applied to tourism product reviews using our corpora. Here, we present the best general performance obtained by doing a sensitivity analysis regarding the most sensitive parameter – the minimum support rule to extract *aspect expressions* as defined in Hu and Liu (2004b). Precision, Recall and *F*-measure were calculated for six different values of this parameter for each task. Then, the best model was chosen using *F*-measure. Table 3 shows the obtained values.

These results show that performance on the *aspect extraction* task is fairly poor in the tourism domain. The algorithm is only capable of extracting almost 30% of the total explicit expressions for hotels and almost 40% for restaurants. Moreover, a high percentage of the extracted expressions do not correspond to real *aspect expressions* for both cases. On the other hand, sentiment classification shows fairly good results, but in this case most of the possible conclusions are difficult to prove because this task was only evaluated for those *aspect expressions* that were extracted. Since these expressions are somewhat the simplest ones, determining the sentiment orientation on them may be easier. Consequently, Precision and Recall could decrease when all *aspect expressions* are considered.

Results also support the properties of tourism product reviews presented in Section 1. These *stories* in which reviewers mention objects that do not correspond to attributes or components of the product may explain the low precision obtained for the explicit *aspect extraction* task in both cases. For instance, in the case of hotels, users commonly refer to objects like time, day and city, which, although relevant for stories, tell nothing about the hotel. Also, nouns and NP sets that do not occur with relative high frequencies will probably need some special treatment in order to be extracted, keeping in mind that many expressions can be used to refer to the same *aspect*. In Hu and Liu (2004b), authors proposed a method to extract these infrequent *aspect expressions* by exploiting their relationships with frequent *opinion words*. Here, this method was not considered since in their case, the extracted infrequent *aspect expressions* only represented an improvement of 15% for Recall, at the cost of decreasing Precision by almost 7%. However, given the poor results that have been obtained, it seems interesting to evaluate how this step would improve or worsen performance in this case. On the other hand, as stated in Hu and Liu (2004b) the reason that probably explains Precision being a little lower than Recall in the task of subjectivity classification is the fact that there are many non-opinionated sentences in tourism product reviews. Since the algorithm labels some of these sentences as opinion sentences because they contain both product *aspect expressions* and some *opinion words*, Precision decreases. Nevertheless, although these sentences may not show strong user opinions toward the product features, they may still be beneficial and useful (Hu & Liu, 2004b).

6.2. Comparing with Liu's approach

Average performance results were computed by simply averaging the results obtained by the best model on both corpora. Table 4 compares these results with the performance obtained by Bing Liu.

From the results in Table 4, an important improvement in relation to the task of extracting subjective sentences can be noticed. In Liu's case, the average recall of opinion sentence extraction is nearly 70%, while the average precision of the same task is 64%.
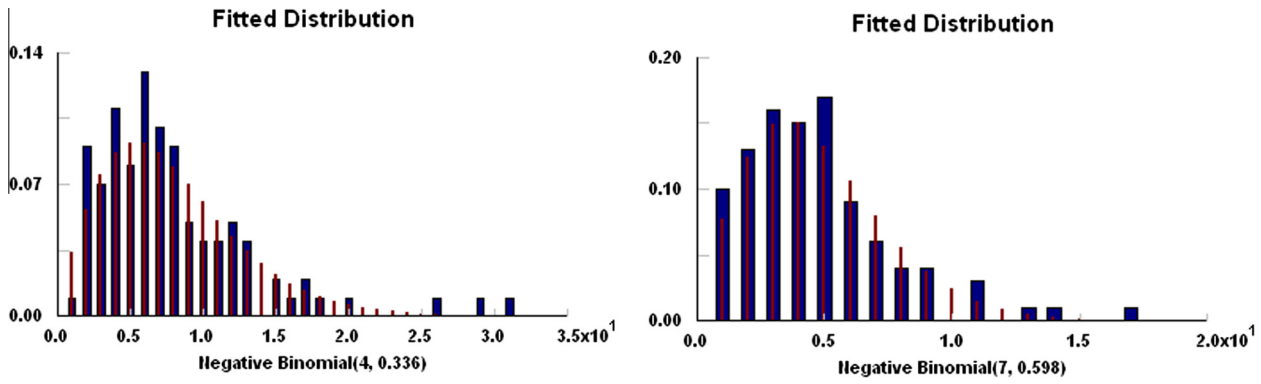

Fig. 3. Charts showing the best-fitting discrete distributions for the number of sentences in each corpus. The chart on the left corresponds to the hotels corpus, the one on the right shows results for the restaurants corpus.
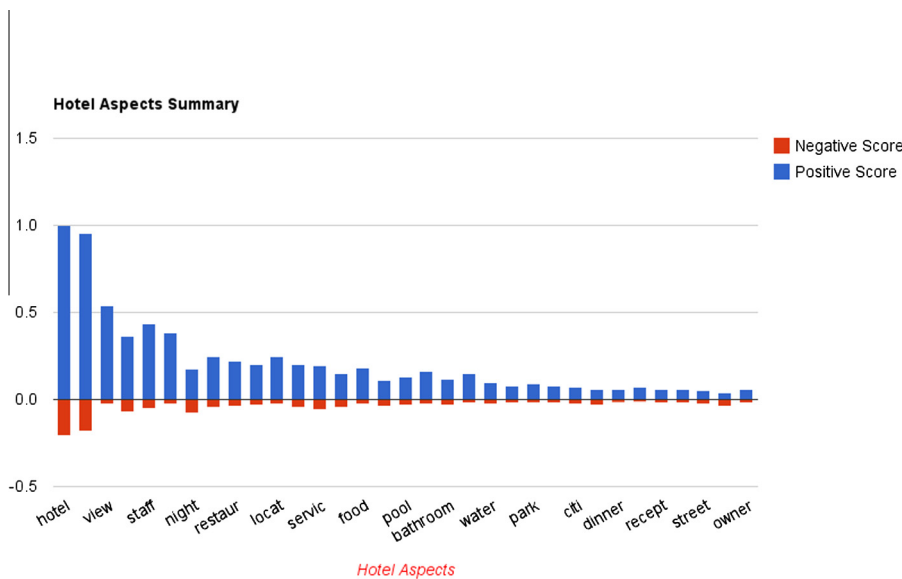
Table 3
Performance results. The bold values Score improvement by using our proposal method.

Corpus	Hotels		Restaurants		Average		
Index	*P* (%)	*R* (%)	*P* (%)	*R* (%)	*P* (%)	*R* (%)	*F*-m (%)
Explicit aspect extraction	33	29	42	37	38	33	36
Subjectivity classification	79	93	81	88	80	**91**	85
Sentiment classification	89	93	91	93	**90**	**93**	**92**

Table 4
Average performance obtained in both corpora, compared with Liu's results. Values not directly given by Liu, obtained using Precision and Recall from Hu and Liu (2004b) and Ding et al. (2008).

Index	Precision		Recall		*F*-measure	
Name	Here (%)	B. Liu (%)	Here (%)	B. Liu (%)	Here (%)	B. Liu (%)
Explicit aspect extraction/P-support pruning Hu and Liu (2004b)	38	79	33	67	36	73*
Subjectivity classification/opinion sentence detection Hu and Liu (2004b)	80	64	91	69	85	67*
Sentiment classification (without 0)/sentiment classification Ding et al. (2008)	90	91	93	90	92	90

Fig. 4. Bar chart for hotels in lake district. Aspects are ordered according to relative importance in a descending manner.

Here, although Precision increased by 10%, the most important improvement is in Recall, in this case 25% higher. On the other hand, sentiment classification shows an improvement, being higher than in Liu's case. Finally, the fact that the aspect extraction task gets poor results, with a decrease of nearly 40% using Liu's unextended approach on tourism product reviews, confirms that
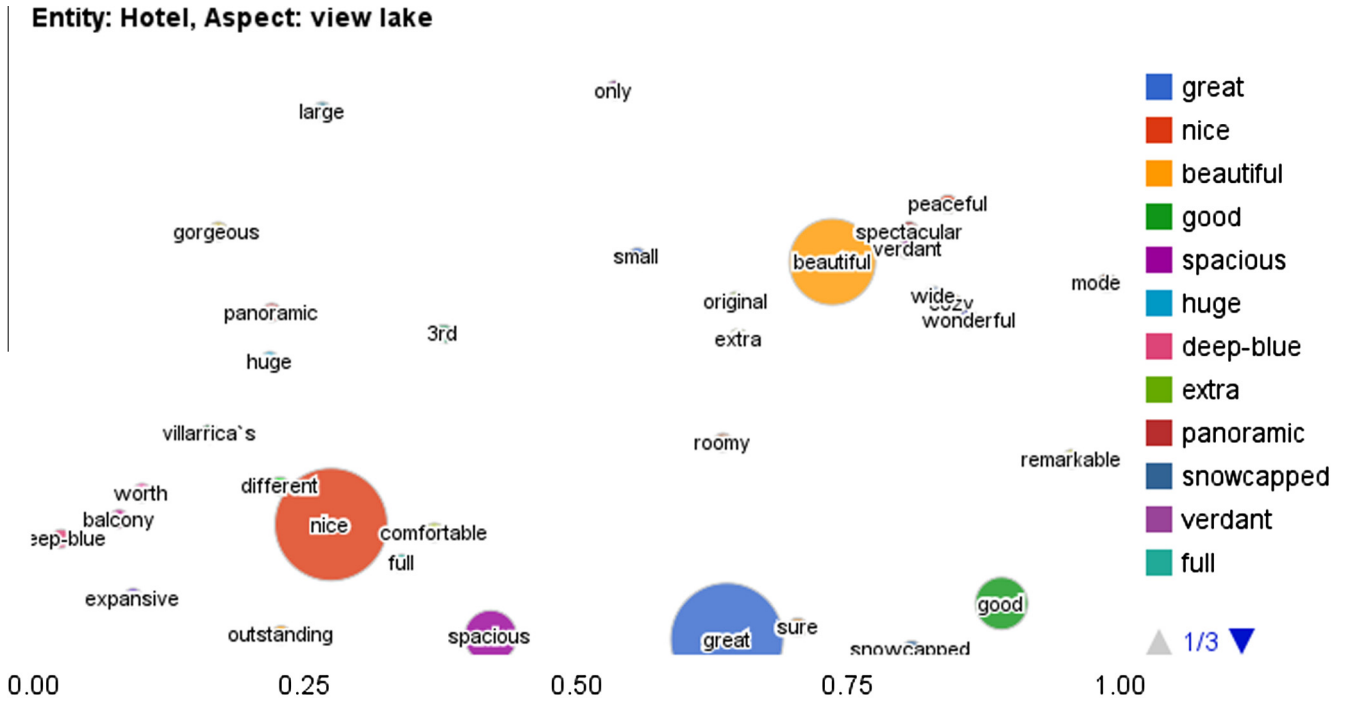
Fig. 5. Proposed bubble charts for the *aspect expression lake view* of hotels.
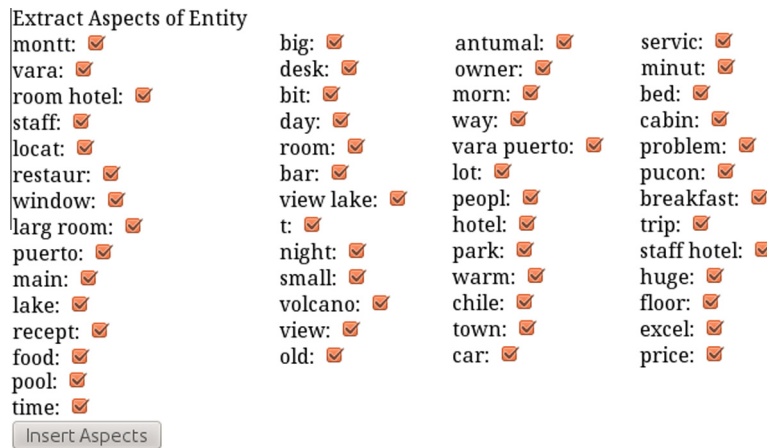

Fig. 6. RVM interface to select *aspect expressions* to be saved in the system.

the features that we found in the domain need to be specially considered in order to get good results.

6.3. Summarization evaluation

Since we intend to show the extracted aspects to users, it is also important to evaluate how the RVM performs. Our Django web-based implementation of the RVM shows aspect-based summaries in which, besides bar charts for each *entity* in the system, a table shows the actual values of the *Positive Score*, *Negative Score* and *Relative Importance* for each *aspect expression*. Fig. 4 shows an example. By clicking the name of each column, the table and the bar chart are sorted according to the clicked column (each click alternates between an ascending or descending sort.)

By clicking one *aspect expression*, the user is redirected to a page showing specific information about it. These pages show the corresponding adjective bubble charts, which are built using the two nearest adjectives. As it is possible to see in Fig. 5, the chart really offers valuable information, indicating that tourists in TripAdvisor tend to describe the lake view using strong positive adjectives, such as nice, great and beautiful. As mentioned before, the RVM also offers users an interface to select the aspects to be saved; Fig. 6 shows how this interface looks. Considering the low obtained performance in the aspect extraction task, this functionality became crucial in this case.

For the evaluation, we first consider the problem of aspect extraction from the perspective of Information Retrieval and measure *precision at k* of the extracted aspects, according to their *Relative Importance*. As shown in Table 5, results prove that this measure ensures that users see aspects with minimum noise. Since the task of sentiment classification has a fairly good performance, we then have empirical evidence that bar charts showing the top *k aspect expressions* deliver accurate and true information to users.

Table 5
Precision at *k* first *aspect expressions*, according to relative importance.

Precision at	Hotels (%)	Restaurants (%)	Average (%)
10	100	90	95
15	73.33	93.3	83
20	75	90	83

Finally, since we wanted to know if the system is able to solve the proposed problem, we interviewed and surveyed a group of 27 Lake District tourism operators, who navigated through the charts as disposed on the website www.patagonialoslagos.cl. In relation to bar charts, 45% of the users completely understood the meaning of each bar without any additional explanation, while for bubble charts most of the users needed help understanding the meaning of the size of each bubble. In general, the charts were difficult to understand mostly for those users that were less familiar with technology or for those that had problems with English. However, results showed that most of the users (almost 80%) considered that the system adds valuable information to their business.

7. Conclusions and future work

In this study, we present a generic design of a tourism opinion mining system that aims to be useful in many industries. The core of our system is an extension of Bing Liu's aspect-based opinion mining technique.

On the one hand, the non-tailored algorithm for *aspect expressions* extraction, based on frequent nouns and NPs appearing in reviews, achieved a poor performance in the tourism domain. This result shows that, in fact, multiple expressions are used to denote the same attribute or component of a tourism product in reviews. Therefore, not only the most frequent words need to be considered when extracting *aspect expressions* in order to achieve a better recall for this task. Likewise, the fact that users tend to *tell stories* when writing reviews about tourism products led to poor precision in the task of extracting *aspect expressions* since in reviews a lot of objects that are not components or attributes of the product are mentioned. In this context, our proposal in sorting the shown aspects according to their *Relative Importance* proved to successfully overcome this problem, ensuring that users see information with minimum noise.

Conversely, the application of NLP rules for determining semantic orientation proved to be very effective for extracted *aspect expressions*, achieving an average Precision and Recall of 90%. Since *aspect expressions* that were extracted only represent a small percentage of the ones that were manually detected, the method needs to be tested for all possible expressions on the topic of tourism in order to give a more conclusive analysis. Also, the good results obtained for the orientation prediction task contribute to ensure that users see noise free information on our system. Thus, the improvement on the performance on both subjectivity and sentiment classification tasks shows that the special rules that were developed by us for the tourism domain are well oriented and allow a better understanding of opinions on that particular domain. However, an important downside of these rules is be the fact that they are not sub-domain sensitive. In the tourism domain, this could represent a major problem since a lot of opinions could imply a positive or negative sentiment depending on the product the opinion is given on.

On a different topic, we realized that the opinion annotation task could easily become very complex. Nevertheless, through the participation of a linguistics expert in the process it was possible to more accurately understand how opinions are given by users and how opinion linguistic corpora should be elaborated.

Documenting any corpora with all the assumptions, rules, techniques or methodologies that were used when generating the input texts or annotating is a key factor to a better understanding for those who may use those corpora. This was a main downside found in Liu's case, considering that in the opinions domain any annotation process will always be a somewhat subjective task.

In conclusion, this work has successfully extended an existing aspect-based mining approach in order to apply it to the tourism domain, particularly, to opinions available on the Web in the manner of tourism products reviews. As a result of the new and more complex NLP-based rules that we developed for both subjective and sentiment classification, our extension is able to perform better than Liu's model, improving both Accuracy and Recall for the mentioned tasks. The effectiveness of these rules shows that the features that we detected on tourism products, such as sentences including multiple mentions of the product or the presence of a high number of sentences containing no opinions, are an accurate characterization of the domain and that they should be considered in future work on the field for a good performance. Likewise, the use of a non-tailored approach for aspect extraction, which led a poor performance, reinforces the importance to consider the special features that exist on this domain.

Finally, in this paper we have also used our proposals to successfully implement a system and tackle the issues in the *Lake District* tourism industry, in the south of Chile. The feedback given by the system users showed that our summarization and visualization charts, which were also proposed as a part of our extension, are easy to understand and give actual insights about opinion, proving how useful and powerful our tool is. Our design and models for aspect-based opinion can be used in many possible applications in the tourism domain. Benefits that may arise entail both tourists and service providers.

7.1. Future work

For future work, the primary objective should be to improve Recall on the task of aspect expression extraction, finding infrequent and implicit *aspect expressions*. Some methods to do so, including the one proposed in Ding et al. (2008), have already been developed in literature but their implementation was here left for the future. Likewise, to improve the Precision of the same task, all the extracted expressions that are not components or attributes of a product need to be filtered. In this context, the use of ontologies as in Cadilhac, Benamara, and Aussenac-Gilles (2010), Zhao and Li (2009) and Vallés Balaguer, Rosso, Locoro, and Mascardi (2010), or other methods of studying relations between words, such as the one proposed in Popescu and Etzioni (2005) or in Bollegala, Matsuo, and Ishizuka (2007), could also be very useful to filter undesired expressions.

On the other hand, we have seen that tourism product reviews contain an important number of sentences that have no opinions. These sentences need to be filtered since they introduce noise to the opinion mining process. This also includes the problem of analyzing context and domain-dependent opinions. New methods to determine subjectivity or sentiment orientation need to be tested on the tourism domain in order to improve the performance of this tasks.

Future work should also tackle the problem of transforming *aspect expressions* into *aspects*. This is a difficult problem yet a crucial feature for any system like ours, because presenting *aspect expressions* to users implies redundancy and makes the analysis more complex. Here, the objective is to build or use ontologies, hierarchies or clusters of *aspect expressions* to make the system become easier to navigate and more intuitive for users.

Finally, a another extension of this work implies working with tourism products reviews written in different languages. Some of

the NLP tasks that are used by our system, including sentence and word tokenizers, are generally machine leaning algorithms that need to be properly trained in order to generate good results. The vast availability of data in English to train these models contrasts with a relative scarcity for other languages. Therefore, there is an immense room for future work on this area.

Acknowledgments

References

Archak, N., Ghose, A., & Ipeirotis, P. (2007). Show me the money!: Deriving the pricing power of product features by mining consumer reviews. In *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 56–65). ACM.

Bollegala, D., Matsuo, Y., Ishizuka, & M. (2007). An integrated approach to measuring semantic similarity between words using information available on the web. In *HLT-NAACL* (pp. 340–347).

Cadilhac, A., Benamara, F., & Aussenac-Gilles, N. (2010). Ontolexical resources for feature based opinion mining: A case-study. In *23rd International conference on computational linguistics* (p. 77).

Cambria, E. (2013). An introduction to concept-level sentiment analysis. In F. Castro, A. Gelbukh, & M. González (Eds.), *Advances in soft computing and its applications. Lecture notes in computer science* (Vol. 8266, pp. 478–483). Berlin Heidelberg: Springer.

Cambria, E., Poria, S., Gelbukh, A., & Kwok, K. (2014). A common-sense based api for concept-level sentiment analysis. *Making Sense of Microposts, 1*(1), 1–2.

Cruz, F. L., Troyano, J. A., Enríquez, F., Ortega, F. J., & Vallejo, C. G. (2013). 'Long autonomy or long delay?' The importance of domain in opinion mining. *Expert Systems with Applications, 40*(8), 3174–3184.

Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing, 27*(4), 293–307.

Ding, X., Liu, B., & Yu, P. (2008). A holistic lexicon-based approach to opinion mining. In *Proceedings of the international conference on Web search and web data mining* (pp. 231–240). ACM.

Dueñas-Fenández, R., & Velásquez, J. D. (2014). LÍuillier, G. Detecting trends on the web: A multidisciplinary approach. *Information Fusion, 20*, 129–135.

Fromkin, V., Rodman, R., & Hyams, N. (2010). *An introduction to language.* Wadsworth Publishing Company.

Havasi, C., Cambria, E., Schuller, B., Liu, B., & Wang, H. (2013a). Knowledge-based approaches to concept-level sentiment analysis. *Intelligent Systems, 28*(2), 12–14.

Havasi, C., Cambria, E., Schuller, B., Liu, B., & Wang, H. (2013b). Statistical approaches to concept-level sentiment analysis. *Intelligent Systems, 28*(3), 6–9.

Hu, M., & Liu, B. (2004a). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168–177). ACM.

Hu, M., & Liu, B. (2004b). Mining opinion features in customer reviews. In *Proceedings of the national conference on artificial intelligence* (pp. 755–760). Menlo Park, CA; Cambridge, MA; London: AAAI Press. MIT Press; 1999.

Hu, M., & Liu, B. (2006). Opinion extraction and summarization on the web. *Proceedings of the national conference on artificial intelligence* (Vol. 21, pp. 1621). Menlo Park, CA; Cambridge, MA; London: AAAI Press. MIT Press; 1999.

Kim, H., Ganesan, K., Sondhi, P., & Zhai, C. (2011). Comprehensive review of opinion summarization. <http://hdl.handle.net/2142/18702>; Accessed 17.11.2013.

Kiss, T., & Strunk, J. (2006). Unsupervised multilingual sentence boundary detection. *Computational Linguistics, 32*(4), 485–525.

Kotler, P. et al. (2001). *Marketing management.* Stuttgart: Verlag Schäffer-Poeschel.

Ku, L.-W., Liang, Y.-T., & Chen, H.-H. (2006). Opinion extraction, summarization and tracking in news and blog corpora. In *AAAI spring symposium: Computational approaches to analyzing weblogs* (pp. 100–107).

Lancaster, K. (1966). A new approach to consumer theory. *The Journal of Political Economy, 74*(2), 132–157.

Liu, B. (2007). *Web data mining: Exploring hyperlinks, contents, and usage data.* Springer Verlag.

Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: Analyzing and comparing opinions on the web. *Proceedings of the 14th international conference on World Wide Web* (pp. 342–351). ACM.

Lu, Y., Zhai, C., & Sundaresan, N. (2009). Rated aspect summarization of short comments. In *Proceedings of the 18th international conference on World Wide Web* (pp. 131–140). ACM.

Marrese-Taylor, E., Rodríguez, O. C., Velásquez, J. D., Ghosh, G., & Banerjee, S. (2013). Web opinion mining and sentimental analysis. In Velásquez J.D., V. Palade, & L. C. Jain (Eds.), *Advanced techniques in web intelligence-2. Studies in computational intelligence* (Vol. 452, pp. 105–126). Berlin Heidelberg: Springer.

Marrese-Taylor, E., Velásquez, J. D., & Bravo-Marquez, F. (2013). Opinionzoom, a modular tool to explore tourism opinions on the web. In *Proceedings of the 2013 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology* (pp. 261–264). IEEE computer society.

Marrese-Taylor, E., Velásquez, J. D., Bravo-Marquez, F., & Matsuo, Y. (2013). Identifying customer preferences about tourism products using an aspect-based opinion mining approach. In *Proceedings of the knowledge-based and intelligent information and engineering systems* (pp. 182–191). KES.

Palakvangsa-Na-Ayudhya, S., Sriarunrungreung, V., Thongprasan, P., & Porcharoen, S. (2011). *Nebular: A sentiment classification system for the tourism business. 2011 eighth international joint conference on computer science and software engineering (JCSSE).* IEEE (pp. 293–298). IEEE.

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

Park, D., & Kim, S. (2009). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications, 7*(4), 399–410.

Popescu, A., & Etzioni, O. (2005). Extracting product features and opinions from reviews. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 339–346). Association for Computational Linguistics.

Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695–729.

Smith, S. L. (1994). The tourism product. *Annals of Tourism Research, 21*(3), 582–595.

Titov, I., & McDonald, R. (2008). A joint model of text and aspect ratings for sentiment summarization. *Urbana, 51*, 61801.

Vallés Balaguer, E., Rosso, P., Locoro, A., & Mascardi, V. (2010). Análisis de opiniones con ontologías. *Polibits, 1*(41), 29–36.

Velásquez, J. D., & González, P. (2010). Expanding the possibilities of deliberation: The use of data mining for strengthening democracy with an application to education reform. *The Information Society, 26*(1), 1–16.

Vicentini, A. (2003). The economy principle in language. Notes and Observations from early modern english grammars. *Mots, Palabras, Words, 3*, 37–57.

Xu, X., Cheng, X., Tan, S., Liu, Y., & Shen, H. (2013). Aspect-level opinion mining of online customer reviews. *Communications, China, 10*(3), 25–41.

Zhao, L., & Li, C. (2009). Ontology based opinion mining for movie reviews. *Knowledge Science, Engineering and Management*, 204–214.

Zhuang, L., Jing, F., & Zhu, X.-Y. (2006). Movie review mining and summarization. In *Proceedings of the 15th ACM international conference on information and knowledge management* (pp. 43–50). ACM.

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing, 74*(2), 133–148.