

Article

Introducing Sentiment Analysis of Textual Reviews in a Multi-Criteria Decision Aid System

Mohammed Jabreel ^{1,2,*} , Najlaa Maarroof ¹, Aida Valls ¹  and Antonio Moreno ¹ 

¹ ITAKA Research Group, Universitat Rovira i Virgili, 43007 Tarragona, Spain; najlamaroof2007@gmail.com (N.M.); aida.valls@urv.cat (A.V.); antonio.moreno@urv.cat (A.M.)
² Department of Computer Science, Hodeidah University, Hodeidah 1821, Yemen
 * Correspondence: mhjabreel@gmail.com

Abstract: Nowadays, most decision processes rely not only on the preferences of the decision maker but also on the public opinions about the possible alternatives. The user preferences have been heavily taken into account in the multi-criteria decision making field. On the other hand, sentiment analysis is the field of natural language processing devoted to the development of systems that are capable of analysing reviews to obtain their polarity. However, there have not been many works up to now that integrate the results of this process with the analysis of the alternatives in a decision support system. SentiRank is a novel system that takes into account both the preferences of the decision maker and the public online reviews about the alternatives to be ranked. A new mechanism to integrate both aspects into the ranking process is proposed in this paper. The sentiments of the reviews with respect to different aspects are added to the decision support system as a set of additional criteria, and the ELECTRE methodology is used to rank the alternatives. The system has been implemented and tested with a restaurant data set. The experimental results confirm the appeal of adding the sentiment information from the reviews to the ranking process.

Keywords: opinion mining; sentiment analysis; aspect-based sentiment analysis; multiple criteria decision aid



Citation: Jabreel, M.; Maarroof, N.; Valls, A.; Moreno, A. Introducing Sentiment Analysis of Textual Reviews in a Multi-Criteria Decision Aid System. *Appl. Sci.* **2021**, *11*, 216. <https://doi.org/10.3390/app11010216>

Received: 17 November 2020

Accepted: 21 December 2020

Published: 28 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The surge of social networks, full of opinions about all possible kinds of objects, has provoked a strong change in decision making. Preferences, prices and other criteria are still very relevant; however, the weight of the opinion of other users is certainly growing [1–3]. Some studies show that a large majority of people reads public reviews on the Web before buying any object [4,5]. Reviews are very important to know the quality of the products and services offered by a company; for example, a restaurant may advertise that it offers a free WiFi service, but the reviews may warn us that the connection quality is very low.

It is arguably important to integrate the information about the preferences of the user with the public opinions on the products in a decision support system. The management of preferences in ranking systems [6,7] and the automated analysis of the sentiments of texts [8] have been heavily studied in their respective fields. However, there are not many current works aiming at the integration of both aspects to rank the decision alternatives.

A multiple criteria decision aid (MCDA) [6] deals with different conflicting criteria, establishing scientific bases to elaborate recommendations according to the needs of decision-makers. Artificial intelligence techniques can be used to elicit and construct a knowledge model for each particular decision-maker.

Sentiment analysis (SA) detects whether the polarity of the opinion of a user in a text is positive or negative. It is commonly applied to online reviews in many different domains [8–10]. The techniques of analysis developed in this field may be applied to a whole document, to a sentence or even to a particular aspect in a sentence (aspect-based sentiment analysis, ABSA) [11]. An aspect refers to a particular characteristic of an object

(e.g., the quality of the food in a restaurant), which may be evaluated in a part of a sentence (e.g., “Although the dinner was very expensive, the food was delicious”).

In this work we analyse how to integrate the public opinions of users about a certain set of products in a standard multi-criteria decision support system. To make this integration it is necessary to identify the aspects appearing in the reviews; to obtain the polarity associated with each aspect in each review; to integrate this information as new criteria to be considered in the MCDM process; and finally, to rank the alternatives considering the user preferences, the knowledge about the values of the alternatives in different criteria and the information about the opinions of the users on the Web.

SentiRank is a system that integrates the opinions of the users on a set of products as a new set of criteria to be taken into account in ELECTRE, the well-known MCDM ranking methodology. This paper focuses especially on the aspect-based sentiment analysis task, which involves two main sub-tasks: aspect detection (detect the aspects of an object mentioned in a sentence) and aspect polarity (identify if the opinion about that aspect of the object is positive or negative).

To summarise, the main contributions of this work are the following:

1. We propose a new sentiment analysis system to obtain the polarity associated with each aspect mentioned in every sentence of a review. First, the relevant aspects of a sentence are detected; after that, a Support Vector Machine (SVM) is employed to calculate the polarity of the opinion of the user for each aspect. The experiments show that this new system has a performance comparable to the one of the current state-of-the-art systems.
2. We integrate the sentiment information about the aspects as new (social) criteria to be considered, alongside the standard criteria of the domain. Then, we apply ELECTRE-III to rank the alternatives.
3. We show how the rankings change when social criteria are considered in a use case with restaurant data.

The rest of the article is structured as follows. Section 2 overviews the state-of-the-art in the fields of SA (specially ABSA) and MCDA. Section 3 describes the development of the aspect-based sentiment analysis system, the architecture of our ranking system and its implementation details. Section 4 presents and discusses the experimental results. Section 5 summarises the article and points out some ideas for future work.

2. Background and Related Works

2.1. Sentiment Analysis

Sentiment analysis is a field of natural language processing that develops mechanisms to analyse the polarity of the opinion expressed in a sentence (or even a part of sentence) automatically. Sentiment polarity is the most well-studied sub-problem of sentiment analysis. It may be conceptualised as a multi-class text classification problem, in which the goal is to determine whether the overall opinion expressed in a given text is positive, negative or even neutral. There are many variants of this task. One is to classify a text according to a rating scale, e.g., 1 = “worst” to 5 = “best”. We may consider different levels of analysis, from the document level to the sentence level. It is also possible to formulate the problem as a regression problem. The goal, in this case, is to predict the intensity of the sentiment that represents better the mental state of the author [12]. The intensity is a real value in the range from 0 (negative) to 1 (positive). Since it is a text classification problem, we can apply any existing supervised learning method, e.g., naive Bayes or Support Vector Machines [13–15].

2.2. Aspect-Based Sentiment Analysis

Classifying text opinions at the document or sentence level is not enough in most applications. With these two levels, we cannot identify the aspects of the object and assign sentiments to them for two main reasons. First, people tend to mention multiple aspects of an entity in their reviews. Figure 1 shows the typical layout of online customer reviews.

Second, a single sentence does not reflect the author's opinion about all aspects of an entity. Hence, we need to perform the analysis at a high level. To this end, we need two steps. First, we need to find all the aspects mentioned in the document. As soon as we get these aspects, we need to decide whether the sentiment expressed towards each aspect is positive or negative. The first step is known as the aspect extraction task, whereas the second step is known as the aspect polarity detection task. These tasks are described in the following sections.

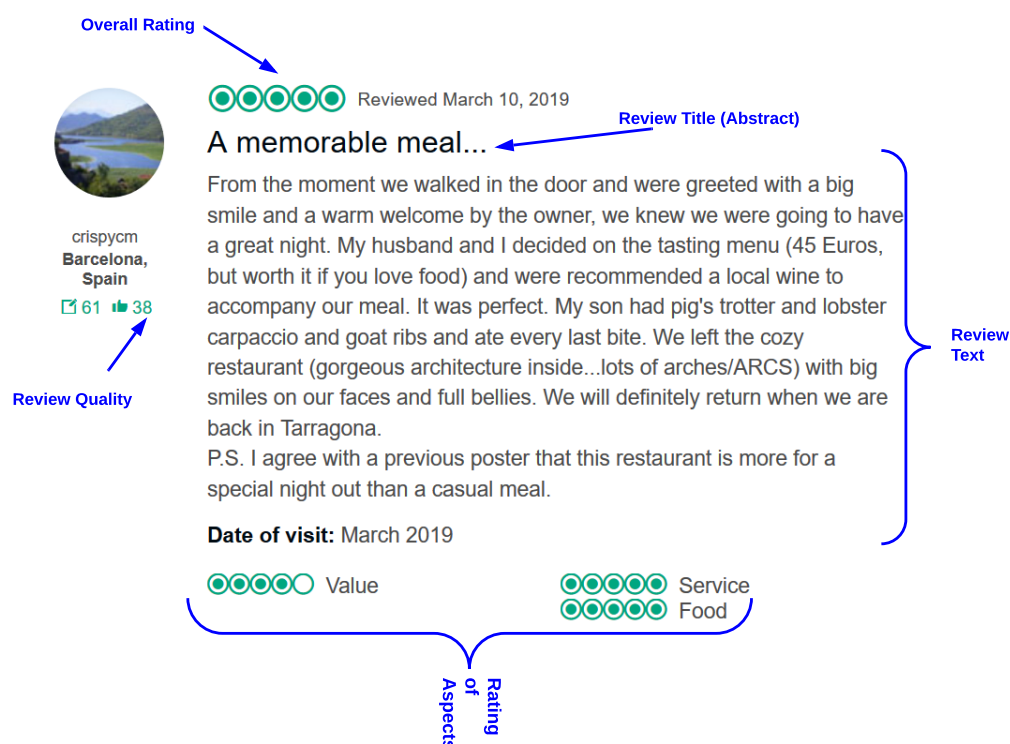


Figure 1. Components of the review of a restaurant.

2.2.1. Aspect Extraction

Aspect extraction aims to extract the set of aspects mentioned in a given text. It is an essential step in opinion mining. The first work on aspect extraction was introduced by Hu and Liu, 2004 [16]. Although they considered explicit and implicit aspects, they only dealt with the former ones. The main idea of their approach was to design a set of rules based on statistical observations. Popescu and Etzioni, 2007 [17] improved this method by assuming that the product class is known a priori. Based on that assumption, they proposed an algorithm that detected whether a noun refers to an aspect by calculating the pointwise mutual information between the product and the noun. Scaffidi et al. [18] improved Hu and Liu's method by assuming that the aspects of a product are more frequent in reviews than in general text. Liu et al., 2015 [19] proposed a novel method to detect aspects by using rules that consider the dependency relations between aspects and the opinion words.

Aspect extraction can be conceptualised as a sequential labelling problem. In this case, the extracted aspects are terms rather than categories. The most popular methods in this context are Hidden Markov Models (HMM) and Conditional Random Fields (CRF). The authors in [20] used a lexicalised HMM to extract opinions along with their explicit aspects. Jakob and Gurevych, 2010 [21], Li et al., 2010 [22] and Choi and Cardie, 2010 [23], employed CRF to extract explicit aspects.

Some studies have employed the multi-label classification approach to extract aspect categories from a text. First, they define a set of aspect categories. Then, they develop a system to map sentences to aspect categories from the defined set. For example, the au-

thors in [24] developed five binary SVM classifiers. Each one specifies whether a sentence belongs to a specific category or not. The NLANGP team [25,26] also modelled aspect category detection as a multi-class classification problem using logistic regression for each category. Sentiu [27] proposed a separate maximum entropy classifier with syntactic features (e.g., words, lemmas). Then, they applied heuristics to the outputs of the classifiers to determine the categories of each sentence. Following this approach, i.e., multi-label classification, we have developed a system to extract aspects from reviews, as explained in Section 3.3.

2.2.2. Aspect Polarity Detection

The goal of aspect polarity detection is the identification of the polarity towards each aspect extracted from a given customer review about a target entity. The aim is to obtain a set of $\{a, p\}$ tuples that represents the user's opinions. The item a is extracted in the aspect extraction step; p can be positive, negative, neutral or conflict. The task can be regarded as a classification problem where the inputs are a sentence, s , and an aspect, a , and the output is the polarity, p . A single sentence may contain opinions about different aspects, with different polarities. The key to solve this problem is to design a set of useful features. Such features need to be able to model the relations between the sentences and the aspects.

In this context, many studies have proposed efficient solutions. For example, the Sentiu team [27] developed a three class polarity classifier by extracting features from the text and the aspect category, including Bag of Words (BoW), lemmas, bigrams, punctuation and polarised words. The ECNU team [28] found that each aspect is related to specific fragments in the corresponding sentence. Based on that idea, they extracted from a given sentence what they called pending words by applying two steps. First, each sentence was divided into several fragments. Then, one or more fragments from the sentence were associated to each aspect. The words in the selected fragments were the pending words. After that, they used four types of features extracted from the pending words to build logistic regression classifiers: linguistic, sentiment lexicon, topic-modelling and word-embedding. The NRC-Can-2014 team [24] trained one multi-class SVM classifier for all aspect categories. The feature set incorporated the information about a certain aspect category. Following these approaches, we have designed our aspect category polarity detection model, as explained in Section 3.4.

2.3. Multi-Criteria Decision Aid

Decision problems [6] are usually complex as they involve the evaluation of objects defined on several conflicting criteria. MCDA methods aim to help users to make a decision when confronted with a set of alternatives [29]. They are widely used in decision problems to find the best alternative. MCDA makes the process more explicit, rational, and efficient. It is a multidisciplinary field, derived from Operations Research, that uses mathematical approaches to deal with complex problems encountered in human activities. Nowadays, it also integrates Artificial Intelligence and economic welfare techniques.

In this work we have used ELECTRE-III, a well-known outranking MCDA method, to compare and order the decision alternatives according to their performance in several features, taking into account the decision maker preferences. The interested reader may look at the details of this method in [30]. For the purposes of this paper, suffice it to say that the user of ELECTRE-III must provide the following parameters for each feature (criterion):

- The weight of the criterion.
- Indifference threshold: if the difference between the values of the criterion in two alternatives A and B is lower than this indifference threshold, then both objects have the same interest for the user and one cannot be preferred to the other.
- Preference threshold: if the difference between the values of the criterion in two alternatives A and B is higher than this preference threshold, then A will be preferred to B in this feature.

- Veto threshold: if the difference between the values of the feature in two alternatives A and B is higher than this threshold, then B will never be preferred to A , even if it has better values in the rest of features.

2.4. Combining MCDA and SA

There are not many previous works that use reviews to rank a set products. Moreover, the existing ones do not take into account the preferences of the user. For example, a method that ranks products considering the public reviews was proposed by Zhang et al. [31]. They first identified subjective and comparative sentences using dynamic programming; then, they employed those sentences to generate a graph of products from which the top ones can be computed.

Peng et al. [32] analysed reviews to identify their main aspects, and they used PROMETHEE to rank the commented products. Zha et al. [33] proposed a similar system. They used an aspect-based sentiment classifier on the reviews and, after that, they ranked the aspects using probabilities taking into account their frequency of appearance in the reviews. Kang et al. [34] calculated the user's preference on a set of options by combining a sentiment analyser and the VIKOR methodology. Liu et al. [35] integrated a lexicon-based sentiment analyser with a representation of the performance of each object based on an intuitionistic fuzzy number. The final ranking was made using PROMETHEE-II. The combination of aspect-based sentiment analysis and intuitionistic fuzzy sets to rank a set of items was also proposed in [36]; in that case, the employed MCDM mechanism was IF-ELECTRE.

Similarly, we propose to use the power of aspect-based sentiment analysis systems and ELECTRE ranking methods to develop a novel ranking system. However, unlike the previous studies, our system is the first that integrates the preferences of the users and the sentiment analysis of the reviews, as shown in the following section.

3. Methodology

Figure 2 shows the architecture of SentiRank, the system that has been designed and developed to rank objects using ELECTRE-III, taking into account the preferences of the user on certain domain attributes and the global opinions of users on the Web about the decision alternatives. It is composed by three main modules. The first one is the customers' reviews unit, which transforms the reviews about alternatives into a matrix of real numbers which we call the social performance table. The second one is the domain analysis unit. It takes as an input the description of the alternatives and the goals of the decision-maker and it converts them to another matrix of real numbers called domain performance table. As soon as we get the output of the two units, we merge them to get the final performance table. Finally, we feed the performance table to the last unit, i.e., the ranking unit, to have the alternatives ranked.

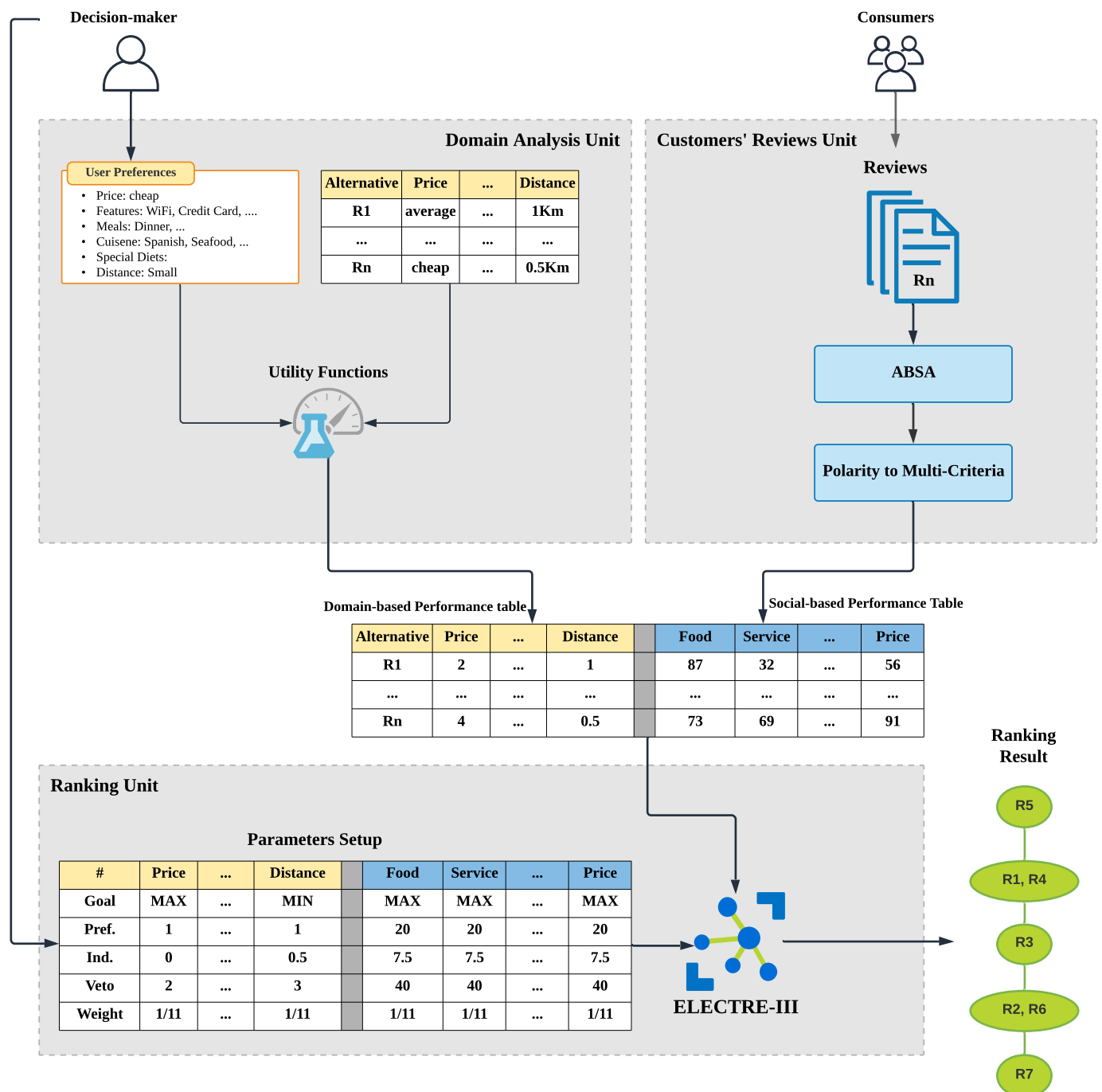


Figure 2. The architecture of SentiRank.

3.1. Customers' Reviews Unit

This module receives a set of reviews about a given set of objects. It analyses the aspects mentioned in each review, it determines their polarity, and then it builds the social performance matrix. In this structure each row represents a particular object, and each column corresponds to the one of the detected aspects. As shown in Figure 3, the first step is to preprocess the text. After that, sentences are separated and tokenised. Then, the aspects mentioned in each sentence are found and their polarity is calculated. All these steps are explained in the following subsections.

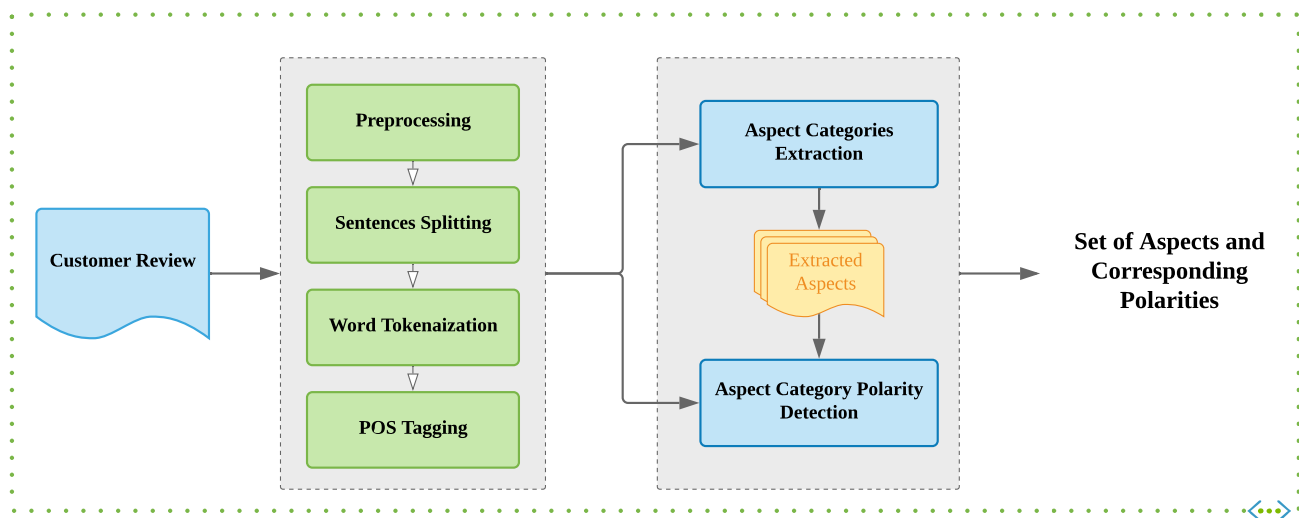


Figure 3. The architecture of the ABSA model.

3.2. Text Preprocessing

Text preprocessing is an essential step for most text analytics problems, especially for sentiment analysis. It is especially important when the text to be analysed comes from reviews on social media, which tend to be written in an informal way. The following text normalisation steps are applied in SentiRank:

- Text is moved to lowercase.
- HTML tags are removed.
- The URLs are replaced by the <URL> symbol.
- The user mentions are replaced by the <USER> symbol.
- The emails are replaced by the <EMAIL> symbol.
- The dates are replaced by the <DATE> symbol.
- The times are replaced by the <TIME> symbol.
- The numbers are replaced by the <NUMBER> symbol.
- The money is replaced by the <MONEY> symbol.

The text was split into sentences, tokenised and tagged with the Part-Of-Speech (POS), using Stanford's tokeniser and POS-tagger.

3.3. Extraction of Aspect Categories

A sentence may contain references to several aspects. Thus, finding these aspects may be conceptualised as a multi-label classification problem [37]. Let $X = \{x_1, x_2, \dots, x_n\}$ be the instance set and $A = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ the aspect set, with dimensions n and m respectively. The supervised multi-label data set D can be defined as follows:

$$D = \{(x_i, \hat{A}_i) | x_i \in X \text{ and } \hat{A}_i \subseteq A \text{ is the set of aspects associated with } x_i\} \quad (1)$$

We use the binary relevance mechanism [37–39] to solve the multi-label classification problem. The main idea is to transform this problem into several binary problems (one for each aspect), that can be solved by independently trained binary classifiers. In order to train the binary classifier for each corresponding aspect α_j , $1 \leq j \leq m$, we first construct the following binary data set:

$$D_j = \{(x_i, \phi(\hat{A}_i, \alpha_j)) | 1 \leq i \leq n\} \quad (2)$$

In this expression, $\phi(\hat{A}_i, \alpha_j)$ is a function that returns 1 if $\alpha_j \in \hat{A}_i$ and 0 otherwise. Figure 4 illustrates the decomposition process.

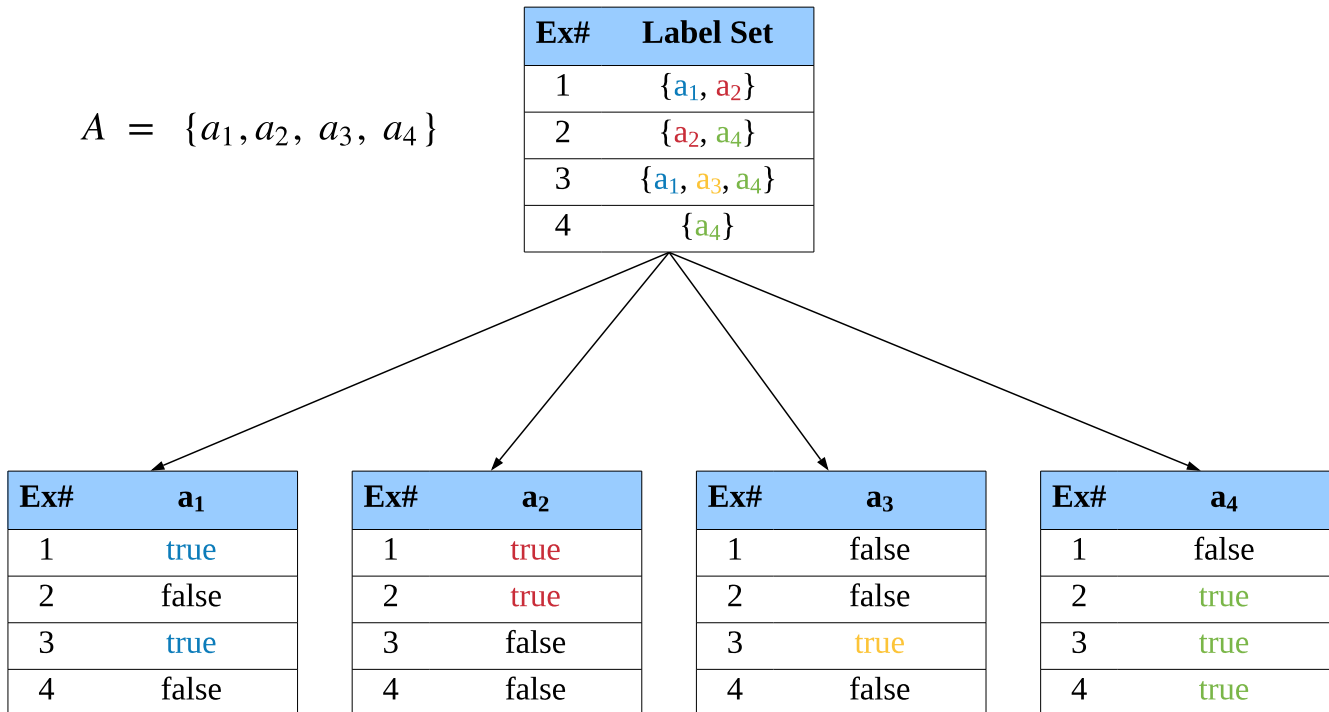


Figure 4. Illustration of the decomposition of a multi-label data set into multiple binary data sets.

After that, we build a SVM binary classifier $h_j : X \rightarrow \{0, 1\}$ from each training set D_j . When the system is given a new sentence x , the aspects that are associated with it are detected by these binary classifiers. Hence, the function $h : X \rightarrow 2^A$ can be defined as follows:

$$h(x) = \{\alpha_j | h_j(x) = 1, 1 \leq j \leq m\} \quad (3)$$

3.4. Calculation of the Aspect Polarity

In this step the system has to calculate whether the opinion of the user about each aspect of each sentence is positive, negative, neutral or “conflict.” A multi-class Support Vector Machine is trained for each aspect category. In the next section it will be seen how the set of features of each SVM is enlarged with specific knowledge about the aspect category.

We define the aspect polarity detection function $f : X \times A \rightarrow \{positive, negative, neutral, conflict\}$ as follows:

$$f(x, \alpha) = \begin{cases} \text{positive} & \text{If } PositionMax(x, \alpha) = 1 \\ \text{negative} & \text{If } PositionMax(x, \alpha) = 2 \\ \text{neutral} & \text{If } PositionMax(x, \alpha) = 3 \\ \text{conflict} & \text{If } PositionMax(x, \alpha) = 4 \end{cases} \quad (4)$$

In this expression $\sigma : X \times A \rightarrow \mathbf{R}^4$ is the classification support function, and the function $PositionMax$ returns the position of the input vector with the highest value. We summarise the aspect extraction and polarity detection steps in Algorithm 1.

Algorithm 1: Aspect extraction and polarity detection algorithm.

```

input : A review  $rev$ .
output: A list of pairs  $L = \{(\alpha_1, p_1), \dots, (\alpha_l, p_l)\}$ , where  $\alpha_k$  is an aspect and  $p_k$  its corresponding polarity.
1  $L \leftarrow \text{Empty}$ ;
2 Split  $rev$  to a set of normalised sentences and store them in  $Sents$ ;
3 for  $x \in Sents$  do
4   Let  $\hat{A}$  be the set of aspects extracted from  $x$  using Equation 3;
5   for  $\alpha \in \hat{A}$  do
6     Let  $p$  be the polarity obtained with Equation (4);
7     Add the pair  $(\alpha, p)$  to  $L$ ;
8   end
9 end

```

3.5. Transforming the Sentiment Information into Social Criteria

In the previous step the system has detected the aspects mentioned in a review and the polarity of the opinion about each of them. In order to integrate this information into a MCDM system, we consider each of the aspects as a social criterion. Thus, we generate a matrix in which each row represents an alternative, and each column is one of the aspects evaluated in the textual reviews. The value for a specific cell is the polarity index [9], which is a number from 0 to 100 that indicates the positiveness of the opinion of the users' reviews about a certain aspect for a given alternative. This index is calculated as follows (P and N are the amounts of positive and negative opinions, respectively):

$$P_{index}(P, N) = \begin{cases} 1 - \frac{N}{P} & \text{if } P > N \\ \frac{P}{N} - 1 & \text{if } P < N \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

The range of P_{index} is $[-1, 1]$, so we transform it to $[0, 100]$ by adding 1 and multiplying the result by 50. Thus, 0 is the lowest satisfaction and 100 represents the highest degree of satisfaction. Algorithm 2 illustrates the transformation steps.

3.6. Domain Analysis Unit

As depicted in Figure 2, this module employs utility functions to transform the values of the domain attributes into numerical values that represent the user's degree of satisfaction, taking into account his/her preferences. The transformation to be made depends on the type of attribute that is being considered. We have considered three kinds of attributes: numerical, categorical and linguistic. In the case of numerical attributes, they do not need any transformation, as they are already numbers. Features of this kind sometimes have to be maximised (e.g., the social features that indicate the opinion gathered from the reviews on some aspects of an object) and sometimes they have to be minimised (e.g., the distance of a restaurant to the city centre). The utility functions needed in the other two cases are described in the following subsections.

3.6.1. Categorical

The utility function used for categorical attributes is the following:

$$g(U; F) = \frac{|U \cap F|}{|U|} \quad (6)$$

In this formula U is the set of values preferred by the user and F is the set of values of the attribute for a particular object. Thus, the utility function computes the percentage of preferred values offered by the object, which the user will desire to maximise.

Example: let us have a categorical criterion (e.g., food types served in a restaurant) and let us assume that the user preferences are $U = \{a, b, c, d, e\}$ and the values provided by an alternative are $F = \{a, d, f, g, h\}$; then, the utility value of this alternative given the preferred values is $g(U; F) = \frac{|\{a, d\}|}{|\{a, b, c, d, e\}|} = \frac{2}{5} = 0.4$.

Algorithm 2: Creation of the social performance table.

input : A set of alternatives O , a set of aspects A and a set of reviews RV .
output: The social performance table, M .

```

1  initialisation;
2   $r \leftarrow |O|$ ;
3   $m \leftarrow |A|$ ;
4   $POS \leftarrow \text{Zeros}[r, m]$ ;
5   $NEG \leftarrow \text{Zeros}[r, m]$ ;
6   $M \leftarrow \text{Zeros}[r, m]$ ;
7   $i \leftarrow 1$ ;
8  while  $i \leq r$  do
9       $RV_i \leftarrow RV[i]$ ;
10     for  $rev \in RV_i$  do
11         let  $L$  be the list of extracted aspects from  $rev$  using Algorithm 1;
12         for  $(\alpha, p) \in L$  do
13             let  $j$  be the index of aspect  $\alpha$ ;
14             if  $p == \text{"positive"}$  then
15                  $POS[i, j] = POS[i, j] + 1$ ;
16             else
17                 if  $p == \text{"negative"}$  then
18                      $NEG[i, j] = NEG[i, j] + 1$ ;
19                 end
20             end
21         end
22     end
23      $i = i + 1$ ;
24 end
25  $i \leftarrow 1$ ;
26 while  $i \leq r$  do
27      $j \leftarrow 1$ ;
28     while  $j \leq m$  do
29          $P \leftarrow POS[i, j]$ ;
30          $N \leftarrow NEG[i, j]$ ;
31          $M[i, j] = P_{\text{index}}(P, N)$ ;
32          $j = j + 1$ ;
33     end
34      $i = i + 1$ ;
35 end

```

3.6.2. Linguistic

The utility function that is employed by the system to transform linguistic values into numbers is the following:

$$g(x; p) = |L| - \text{abs}(\text{pos}(x) - \text{pos}(p)) \quad (7)$$

In this formula L is the set of linguistic terms used in that attribute, the function pos returns the position of a particular value in this set, p is the term preferred by the user and x is the term appearing in a certain alternative. The utility function returns a number

between 1 and the number of terms; the higher this number, the higher the satisfaction of the user with that alternative.

Example: Let us have the price feature represented as a linguistic variable. Let $L = \{very\ low, low, high, very\ high\}$ be the linguistic values. Here, $pos(low)$ is 2. Assuming that the user prefers $p = low$ and the alternative value is $x = very\ high$, then the utility value of this alternative given the user preference is $g(very\ high; low) = 4 - abs(4 - 2) = 4 - 2 = 2$.

3.7. Ranking Unit

This module receives the performance table (which integrates the domain attributes and the social attributes derived from the sentiment analysis of the reviews) and ranks the alternatives using ELECTRE-III. In order to use this methodology, it is necessary to know the range of values of each attribute, whether an attribute has to be maximised or minimised to satisfy the user, the weight to be given to each attribute and the ELECTRE thresholds of each criterion (preference, indifference and veto). ELECTRE-III compares each pair of alternatives and creates the outranking relations between them; after that, the exploitation step is done to compute the final ranking of the alternatives.

4. Experiments and Results

A set of restaurants from the city of Tarragona was chosen as a use case to test the system, considering users with different preferences. This section is structured as follows. First, we describe in Section 4.1 the data and the criteria used in this study. Section 4.2 explains how we built the sentiment analysis models. We show the features that have been used to extract the aspects and to evaluate the polarity of the users' opinions of them, and the values of the parameters of the employed classifiers. We also show and discuss the performances of the developed models with a publicly available benchmark on the restaurant domain presented in SemEval-2014 [40]. In Section 4.5, we show the performance of SentiRank by varying the preferences of the user and analysing the obtained results.

4.1. Data

Table 1 shows all the criteria that define the peculiarities of each restaurant. The domain criteria include the information of the restaurants provided in the TripAdvisor website. The sentiments of the reviews were analysed with respect to five aspect categories, as described in the ABSA task of SemEval-2014: food, price, service, ambience and anecdotes [40]. The polarities of the reviews on these categories constitute the values of the social criteria.

The models to detect the aspects and determine the polarity of the user's opinion of them were built and tested using the training and testing sets provided in the ABSA task of the SemEval-2014 competition [40]. They had 3041 and 800 sentences in English, respectively. Each sentence was labelled with a set of pairs (aspect, polarity). The number of opinions for each aspect and each polarity is shown in Table 2. Most of the sentences contain positive opinions, and the aspect that received the most opinions was the quality of the food.

Table 1. Domain and social criteria used in the case study.

Criterion	Attribute Type	Domain	Goal
A-Domain			
Price	Linguistic	{Very Low, Low, High, Very High}	MAX
Cuisines	Categorical	Italian, Chinese, Grill, Bar, ...	MAX
Special Diets	Categorical	Vegetarian Friendly, ...	MAX
Meals	Categorical	Lunch, Drinks, ...	MAX
Features	Categorical	Private Dining, Seating, WiFi, ...	MAX
Distance-to-city-centre	Numerical	[0, ∞]	MIN
B-Social			
Food	Numerical	[0, 100]	MAX
Price	Numerical	[0, 100]	MAX
Service	Numerical	[0, 100]	MAX
Ambience	Numerical	[0, 100]	MAX
Anecdotes	Numerical	[0, 100]	MAX

Table 2. Aspect categories distribution per sentiment class.

Aspect	Positive		Negative		Conflict		Neutral		Total	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
FOOD	867	302	209	69	66	16	90	31	1232	418
PRICE	179	51	115	28	17	3	10	1	321	83
SERVICE	324	101	218	63	35	5	20	3	597	172
AMBIENCE	263	76	98	21	47	13	23	8	431	118
ANECD	546	127	199	41	30	15	357	51	1132	234
Total	2179	657	839	159	163	52	500	94	3713	1025

4.2. Sentiment Analysis Model: Training Setup

We describe in this subsection the training setup of the aspect extraction and the aspect polarity detection models. We used LIBLINEAR (<https://www.csie.ntu.edu.tw/~cjlin/liblinear/>) to build all the linear classifiers. First, we define all the features used in our system, and then we show the configuration of features for each model. After that, we present the evaluation metrics and the performances of the models, and finally we discuss the results on the Tarragona case study.

4.2.1. Features

The features used in our system are the following:

- **Word n-grams:** appearances of 1, 2, ..., n consecutive words.
- **Character n-grams:** appearances of 3, 4, ..., n consecutive characters.
- **Part-Of-Speech tags:** appearances of each kind of POS.
- **Lexicon-Based features:** A sentiment lexicon associates to a word a number that represents its degree of positiveness/negativity. For each word w and polarity value $score(w; l)$ in a given lexicon l , we calculate the following features: (1) the total score $\sum_w score(w; l)$; (2) the highest value $\max_w score(w; l)$; (3) the number of positive words in the sentence; (4) the lowest value $\min_w score(w; l)$; (5) the number of negative words

in the sentence, and (6) the value of the last word in the sentence. The lexicons used in this work are the following:

- Yelp Restaurant Sentiment Lexicon (Yelp-Res): this lexicon, which contains almost 40,000 unigrams, was created automatically from restaurant reviews contained in Yelp [41] (http://www.yelp.com/dataset_challenge).
- NRC Hashtag Sentiment Lexicon (NRC-Hashtag): This lexicon was constructed from a pseudo-labelled corpus of tweets [41]. It contains almost 40,000 unigrams and 180,000 bigrams.
- NRC Yelp Word-Aspect Association (WA-Lexicon): It contains lexicons of words associated with the five aspect categories shown in Table 2, created from over 180,000 reviews in Yelp. Thus, unlike the other lexicons, this one is centred in the aspect categories. The value that represents the relationship between each word and each category was calculated using the PMI metric [24].
- **Word Clusters:** The CMU POS-tagger (<http://www.cs.cmu.edu/~ark/TweetNLP/clusters/50mpaths2>) employed Brown clustering to divide the English words appearing in over 50 million tweets in 1000 clusters. We used the total occurrences of words from each of the clusters as features.

4.2.2. Configurations

Table 3 shows the features employed to build the aspect detection module. Their values were obtained using 5-fold cross-validation.

Table 3. The features used to detect each aspect category.

	Word-Ngram	Char-Ngram	POS	Clusters	WA-Lexicon
Food	✓, N = 3		✓	✓	✓
Price	✓, N = 4		✓	✓	
Service	✓, N = 1	✓, N = 6		✓	
Ambience	✓, N = 3		✓	✓	✓
Anecdotes	✓, N = 3		✓	✓	✓

As suggested in [24], a domain adaptation mechanism was used to enlarge the features used to train this module. Given an aspect category a , two copies were made of each word x . They are called $x\#general$ and $x\#a$. The former refers to the general use of the word, whereas the latter represents the usage of the word when referring to a particular aspect. In that way, the classifier may learn that a certain word is positive in general (e.g., “splendid”), and it also may learn that a word is positive in a particular context (e.g., “tasty” is a very good word when referring to the quality of food).

As described in Section 3.2, first the sentences of the reviews are separated into words and POS-tagged. After that, the system calculates all the features mentioned in the previous section, except the n-grams based on characters: word 3-grams (both general and aspect-based), Part-Of-Speech tags and the cluster and lexicon-based features (with the three lexicons mentioned above).

4.3. Evaluation Metrics

To evaluate the module of aspect detection, the standard precision (P), recall (R) and F1 measures were used (S is the set of aspects calculated by SentiRank, and G is the set of correct aspects of each sentence).

$$P = \frac{|S \cap G|}{|S|} \quad (8)$$

$$R = \frac{|S \cap G|}{|G|} \quad (9)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (10)$$

The correctness of the polarity of the opinion with respect to an aspect is measured by the accuracy (correct polarity labels associated with an aspect by the system divided by the number of correct labels).

4.4. Results

Table 4 shows the performance of the aspect detection module for each category. In general, the model shows a very good *F1* value in all the categories. The model gives the best performance in the food and price categories, whereas the worst value appears in the anecdotes category.

Table 4. Evaluation of the aspect detection module.

Aspect	P	R	F1	Accuracy
Food	94.43	89.23	91.76	91.63
Service	93.75	87.21	90.36	96
Price	98.61	85.54	91.61	98.25
Ambience	86.73	83.05	84.85	95.63
Anecdotes	80	75.21	77.53	87.25
Average	90.70	84.04	87.22	93.75

SentiRank was also compared with the top three systems in the ABSA task in SemEval-2014 [40], as shown in Tables 5 and 6. The best values are shown in bold. The results for SentiRank are quite close to the ones of the top system, NRC-Can-2014 [24], in both tasks.

Table 5. Evaluation of the aspect detection module.

System	P	R	F1
NRC-Can-2014 [24]	91.04	86.24	88.58
UNITOR [42]	84.98	85.56	85.26
XRCE [43]	83.23	81.36	82.28
SentiRank	90.41	84.68	87.45

Table 6. Evaluation of the polarity calculation module.

System	Accuracy
NRC-Can-2014 [28]	82.92
UNITOR [42]	76.29
XRCE [43]	78.14
SentiRank	82.44

To get more insights into the system's performance, we analysed the confusion matrices shown in Figure 5. The figure on the top left is the confusion matrix of the overall score, when we ignore the category and consider only the polarities. The other five figures show the confusion matrices at the category level of analysis. In general the system's performance is quite good, as 82.8% instances were correctly classified. The best values

correspond to the positive and negative classes, whereas the worst ones appear in the conflict and neutral classes, where most samples were misclassified as positive. To understand the performance of the system for each category, we analysed the confusion matrix for each of them. The best performance appeared in the food, service and price categories. The ambience category showed the worst performance, followed by the “Misc” category. Such a finding can be attributed to the low number of training examples for these two categories. As the positive class had the highest number of training samples in all the categories, it showed a remarkable performance in all of them. However, on the other hand, this superiority of the positive examples caused a bias in the classifications and led to a low performance in other categories.

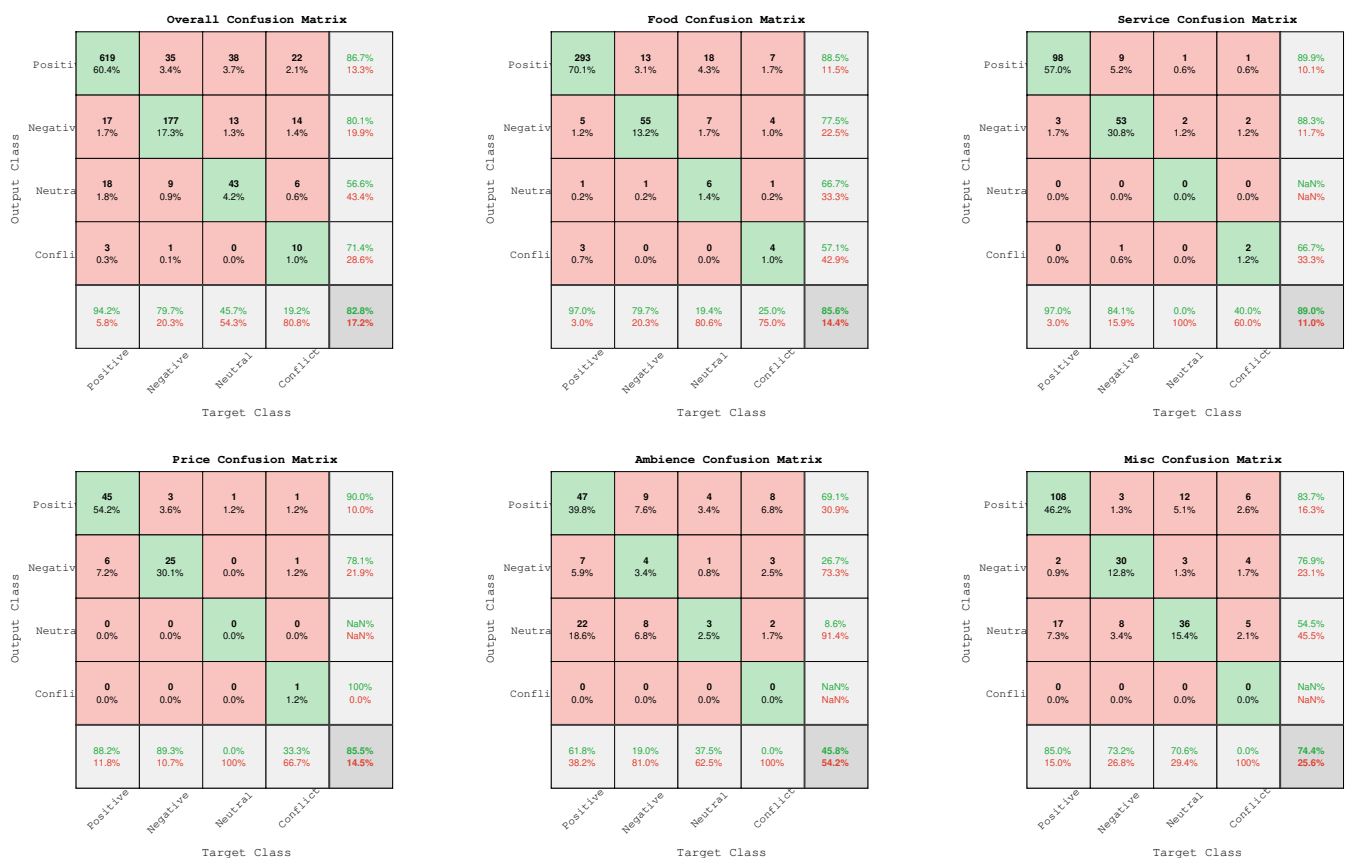


Figure 5. Confusion matrices.

4.5. Case Study

The information available in TripAdvisor about 23 local restaurants was collected. As it may be seen in Figure 6, there is information about the number of reviews, the distance to the city centre, the range of prices, the types of food and the catering of special culinary requirements. There is also a rating for each domain attribute and a global rating of the restaurant. The ratings for the 23 alternatives are shown in Table 7. Interested readers may find the remaining details in TripAdvisor.

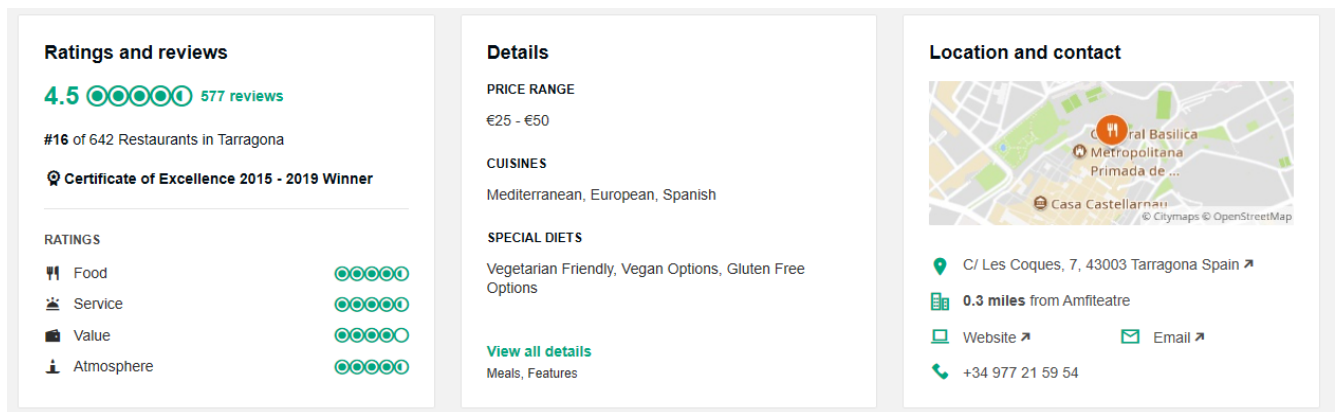


Figure 6. Restaurant details in TripAdvisor.

Table 7. The alternatives and the ratings based on TripAdvisor.

Id	Restaurant Name	Overall	Food	Service	Price	Amb.
1	AQ	4.5	4.5	4.5	4	4.5
2	ELIAN Cafe Restaurant	4.5	4.5	4.5	4.5	4.5
3	La Caleta	4.5	4.5	4.5	4	4.5
4	Tarakon	4.5	4.5	4.5	4.5	4
5	Sadoll Restaurant	4.5	4.5	4.5	4.5	4.5
6	El Taller	4.5	4.5	4.5	4	4
7	El Trull	4.5	4.5	4.5	4.5	4.5
8	Les Coques	4	4	4	4	4
9	Mas Rosello	4	4.5	4.5	4	4.5
10	Les Fonts de Can Sala	4.5	4.5	4.5	4	4.5
11	La Xarxa	4.5	4.5	4.5	4.5	4
12	El Encuentro	4.5	4	4	4.5	3.5
13	La Capital	4	4	4.5	4	3.5
14	Barquet	4.5	4.5	4	4	4
15	Arcs Restaurant	4.5	4.5	4.5	4	4.5
16	Restaurante Ca L Eulalia	4.5	4.5	4	3.5	3.5
17	Restaurante Club Nautico Salou	4.5	4.5	4.5	4	4.5
18	Octopussy	3.5	3.5	3.5	3.5	3.5
19	Palermo 1962 S.c.p.	4	4	4	4	3
20	The Cotton Club Restaurant & Cocktails	4.5	4.5	4.5	4.5	3.5
21	Lizarran Parc Central	3	3.5	3.5	3	3
22	Buffalo Grill	3	3.5	3.5	3	4
23	Indian Restaurant Mirchi Tarragona	3	2	2	3	3

We analysed the correlation between the TripAdvisor ratings and the results of the ABSA module as follows. First, we calculated the polarity index of each pair (alternative, aspect) given a set of reviews as described in Algorithm 2. Then, the overall polarity index of an alternative was obtained by taking the average of all the aspects. Finally, we converted the polarity index into a rating value from 0 to 5 by dividing the polarity index value by 20. The obtained correlations were 0.93 (food), 0.83 (service), 0.55 (price), 0.73 (ambience) and

0.8 (overall rating). Thus, the correlation between the overall rating provided by our system and the overall rating provided by TripAdvisor is strong. This shows the robustness of our sentiment analysis system and indicates that it is applicable to the restaurant domain. The low correlation in the “price” aspect may be attributed to the fact that people tend to give their opinion towards the quality of the service rather than the price, as we found that the “price” aspect has the lowest number of reviews. The “ambience” aspect shows an acceptable correlation with a value of 0.73; however, it is not quite strong. This can be attributed to the low performance of our system with respect to this aspect as shown in Table 4.

As the results discussed above show the effectiveness of our sentiment analysis system, now we can use it to combine the preferences of the users with the analysis of the reviews to rank the restaurants. In the next subsection we will show the results of SentiRank considering two different kinds of users, and examining the influence of the addition of the social criteria derived from the sentiment analysis.

4.5.1. Example 1

In this first example we are going to consider a user with the following preferences:

- **Price:** low.
- **Cuisine:** bar, seafood, Spanish.
- **Special Diets:** no gluten, vegetarian.
- **Meals:** brunch, cocktails, lunch.
- **Features:** parking, private dining, high chairs, handicap accessible, free WiFi, credit cards accepted.
- **Distance:** the distance to the city centre must be as small as possible.

The user has some flexibility in the type of food and the general services of the restaurant, but it is not as tolerant with regard to the times in which meals are served and the diet—special requirements. The values of the ELECTRE thresholds that represent this knowledge are shown in Table 8. All the social criteria are treated in the same way.

Table 8. The values of the ELECTRE thresholds for all criteria.

Criterion	Indifference	Preference	Veto
A-Domain			
Price	0	1	2
Cuisine	0	0.4	No Veto
Special Diets	0	0	0.5
Meals	0	0	0.5
Features	0	0.4	No Veto
Distance	0.5	1	3
B-Social			
Food	7.5	20	40
Price	7.5	20	40
Service	7.5	20	40
Ambience	7.5	20	40
Anecdotes	7.5	20	40

Figures 7 and 8 show the domain and social performance tables respectively. Figure 9 shows the ranking results using only the standard criteria of the domain (left) and also taking into account the social criteria (right). As expected, the addition of the social

criteria derived from the analysis of the textual reviews produce changes in the ordering of the restaurants.

Seq	Restaurant	Price	Cuisines	Special Diets	Meals	Features	Distance
1	AQ	2	0.33	1	0.33	0.67	2.6
2	ELIAN Cafe	3.5	0.33	1	0.33	0.33	0.5
3	La Caleta	2	0.33	1	0.33	0.89	3.5
4	Tarakon	3	0.33	0.5	0.33	0.33	1.9
5	Sadoll	3.5	0.33	1	0.33	0.67	2.1
6	El Taller	3.5	0.33	1	0.33	0.78	1.4
7	El Trull	3.5	0.33	1	0.33	0.67	7.6
8	Les Coques	2	0.33	1	0.33	0.56	2.3
9	Mas Rosello	2	0.33	1	0.33	0.67	6.1
10	Les Fonts de Can Sala	2	0.00	1	0.33	1.00	3
11	La Xarxa	3.5	0.67	0.5	0.33	0.78	1.6
12	El Encuentro	3	0.33	0	0.33	0.67	0.9
13	La Capital	3	0.67	0.5	0.00	0.78	1
14	Barquet	3.5	0.67	1	0.33	0.56	1.1
15	Arcs	2	0.33	1	0.33	0.67	2.5
16	Ca L Eulalia	2	0.67	0.5	0.33	0.56	2.3
17	Club Nautico Salou	3.5	0.67	0	0.33	0.89	14.3
18	Octopussy	3.5	0.67	0.5	0.67	0.89	2
19	Palermo 1962 S.c.p.	3.5	0.67	0.5	0.33	0.44	11.2
20	The Cotton Club & Cocktails	3.5	0.33	1	0.67	0.78	11.3
21	Lizarran Parc Central	3.5	0.33	0	0.33	0.56	1.3
22	Buffalo Grill	3.5	0.00	0	0.33	0.33	5.7
23	Indian Mirchi	3.5	0.00	0	0.00	0.11	0.75

Figure 7. Example 1: domain performance table. The colour indicates the degree of satisfaction, green being the worst and red the best.

Seq	Restaurant	Food	Service	Price	Ambience	Anecdotes
1	AQ	95.51	91.67	38.89	91.30	93.75
2	ELIAN Cafe	98.75	93.33	100.00	83.33	96.88
3	La Caleta	97.22	76.92	100.00	100.00	96.15
4	Tarakon	96.43	100.00	100.00	75.00	90.38
5	Sadoll	95.28	95.00	95.00	96.15	96.30
6	El Taller	97.83	91.67	83.33	100.00	90.91
7	El Trull	97.22	76.92	100.00	100.00	96.15
8	Les Coques	92.47	68.42	33.33	96.15	91.67
9	Mas Rosello	96.15	100.00	87.50	97.22	100.00
10	Les Fonts de Can Sala	97.22	100.00	100.00	100.00	75.00
11	La Xarxa	100.00	80.00	75.00	75.00	97.37
12	El Encuentro	92.00	75.00	87.50	66.67	83.33
13	La Capital	81.25	87.50	100.00	83.33	73.08
14	Barquet	92.50	82.50	81.25	80.77	82.00
15	Arcs	95.08	91.59	75.00	94.26	95.08
16	Ca L Eulalia	96.88	84.62	35.71	78.57	92.31
17	Club Nautico Salou	91.67	85.71	83.33	96.15	100.00
18	Octopussy	50.00	14.71	50.00	16.67	70.00
19	Palermo 1962 S.c.p.	93.65	92.31	70.00	77.27	90.48
20	The Cotton Club & Cocktails	87.01	86.50	78.89	92.22	87.59
21	Lizarran Parc Central	50.00	50.00	50.00	50.00	50.00
22	Buffalo Grill	78.95	81.25	50.00	50.00	100.00
23	Indian Mirchi	20.00	10.00	50.00	0.00	25.00

Figure 8. Example 1: social performance table. Colours have the same meanings as in the previous table.

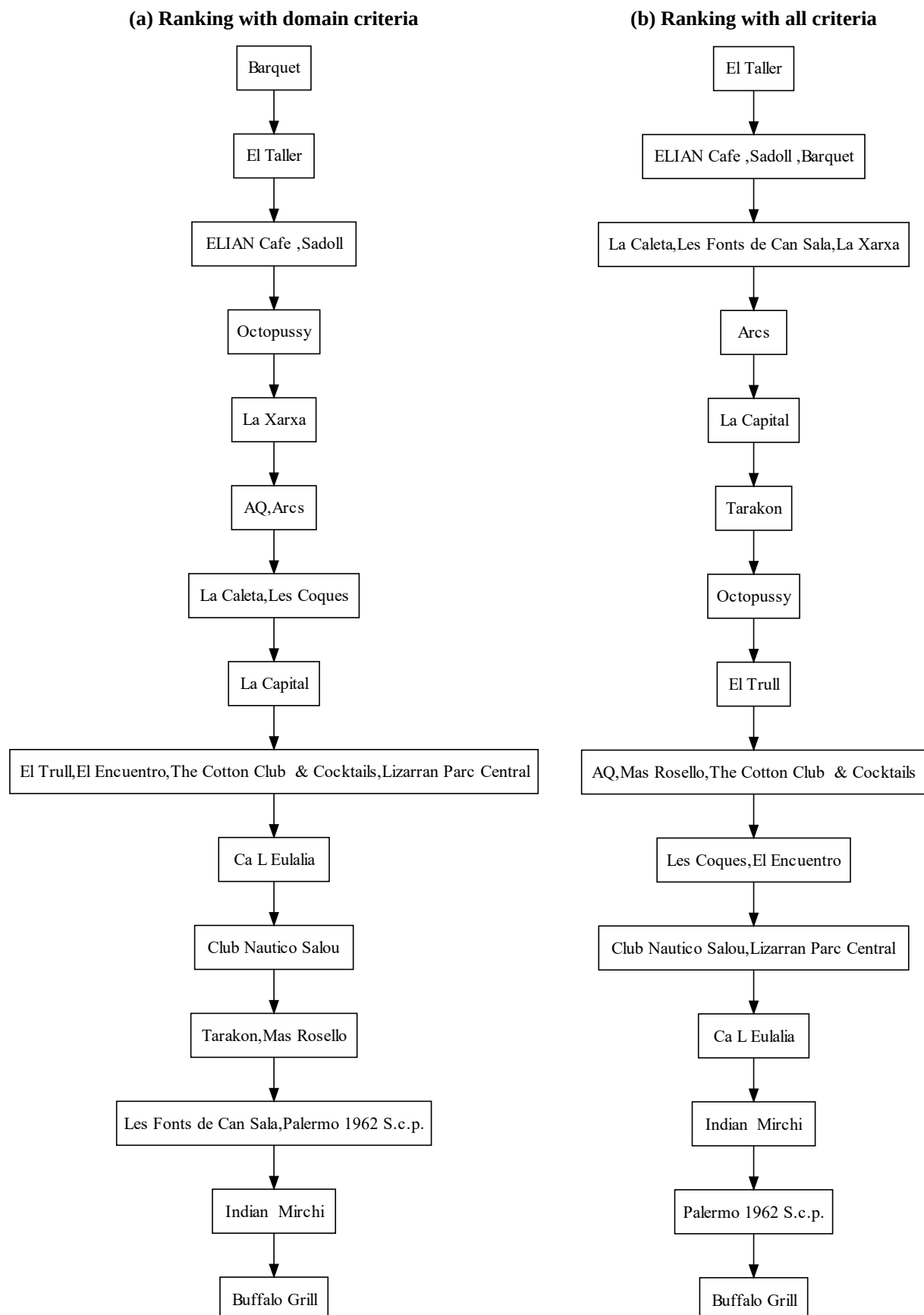


Figure 9. Profile 1: ranking results.

The difference in the rankings reflects the importance of taking into account the opinions of the users expressed in the reviews. In the case of domain criteria, as it may be seen in Table 7, most of the criteria have similar values (especially those for which the user has a stronger preference). The attribute in which there is a greater difference between the restaurants is the distance to the city centre.

Several interesting facts may be noted when considering social criteria:

- The restaurant “El Taller” has better reviews than “Barquet” in all aspects, as shown in Figure 8, which allow it to obtain the first position when these opinions are taken into consideration. For the same reason, “Sadoll” and “ELIAN Cafe” reach the same level as “Barquet”.
- The restaurant “Les Fonts de Can Sala” improves very heavily its position in the ranking when the results of the sentiment analysis are considered. In Figure 8 it may be seen that it is the only option that has perfect reviews in three of the five social criteria. “Tarakon” is another example of a place with very good reviews in four of the five aspects that make it improve its position in the ranking.
- The three restaurants at the bottom are not able to improve their positions, as the customers’ opinions about them are not very good.

In summary, this example shows the clear influence of the social criteria on the ranking process.

4.5.2. Example 2

In this second example the following user preferences are considered:

- **Price:** very high.
- **Cuisine:** grill, barbecue, bar.
- **Special Diets:** no gluten, vegetarian.
- **Meals:** dinner, drinks, late night.
- **Features:** private dining, serves alcohol, table service, free WiFi, credit cards, outdoor seating, parking, full bar.
- **Distance:** restaurants in the city centre are preferred.

We keep the same weight for all criteria and the same ELECTRE thresholds as in the previous example. The utility values of the domain criteria are shown in Figure 10 and the ranking results of the two cases (without and with the social-based criteria) are shown in Figure 11. As we can see, there are clear differences in the two ranking graphs. For example, the alternative “La Caleta” changes from the third to the first position. On the other hand, the alternative “Les Coques” is downgraded from position 1 to position 8 and “AQ” is downgraded from position 2 to position 7. The analysis of the domain performance table for this example, depicted in Figure 10, shows that the more relevant criteria are price, distance to the city centre and features, in this order.

Changes similar to those reported in the first example may also be noted here. For example, the restaurant “Les Coques”, which was the best in terms of domain criteria, goes down to the middle of the ranking when users’ opinions are considered.

Given all that, we can conclude that the incorporation of the social-based criteria adds more constraints to the ranking process. However, we believe that the integration of the users’ reviews with the decision-maker preferences needs more investigation to improve the results. The next section discusses this point in more detail.

Seq	Restaurant	Price	Cuisine	Special Diets	Meals	Features	Distance
1	AQ	4.00	0.00	1.00	0.33	0.63	2.60
2	ELIAN Cafe	2.50	0.00	1.00	0.33	0.50	0.50
3	La Caleta	4.00	0.00	1.00	0.33	1.00	3.50
4	Tarakon	1.00	0.00	0.50	0.33	0.25	1.90
5	Sadoll	2.50	0.00	1.00	0.33	0.50	2.10
6	El Taller	2.50	0.00	1.00	0.33	0.63	1.40
7	El Trull	2.50	0.67	1.00	0.33	0.75	7.60
8	Les Coques	4.00	0.00	1.00	0.67	0.50	2.30
9	Mas Rosello	4.00	0.00	1.00	0.67	0.63	6.10
10	Les Fonts de Can Sala	4.00	0.00	1.00	0.00	1.00	3.00
11	La Xarxa	2.50	0.00	0.50	0.33	0.63	1.60
12	El Encuentro	1.00	0.00	0.00	0.33	0.63	0.90
13	La Capital	1.00	0.33	0.50	0.00	0.75	1.00
14	Barquet	2.50	0.00	1.00	0.67	0.38	1.10
15	Arcs	4.00	0.00	1.00	0.33	0.50	2.50
16	Ca L Eulalia	4.00	0.00	0.50	0.33	0.50	2.30
17	Club Nautico Salou	2.50	0.00	0.00	0.33	0.88	14.30
18	Octopussy	2.50	0.00	0.50	0.67	0.88	2.00
19	Palermo 1962 S.c.p.	2.50	0.00	0.50	0.33	0.38	11.20
20	The Cotton Club & Cocktails	2.50	0.00	1.00	1.00	0.75	11.30
21	Lizarran Parc Central	2.50	0.00	0.00	0.33	0.38	1.30
22	Buffalo Grill	2.50	0.33	0.00	0.33	0.13	5.70
23	Indian Mirchi	2.50	0.00	0.00	0.00	0.13	0.75

Figure 10. Profile 2: the domain performance-table. The colour indicates the degree of satisfaction, and colours range from green, the worst, to red, the best.

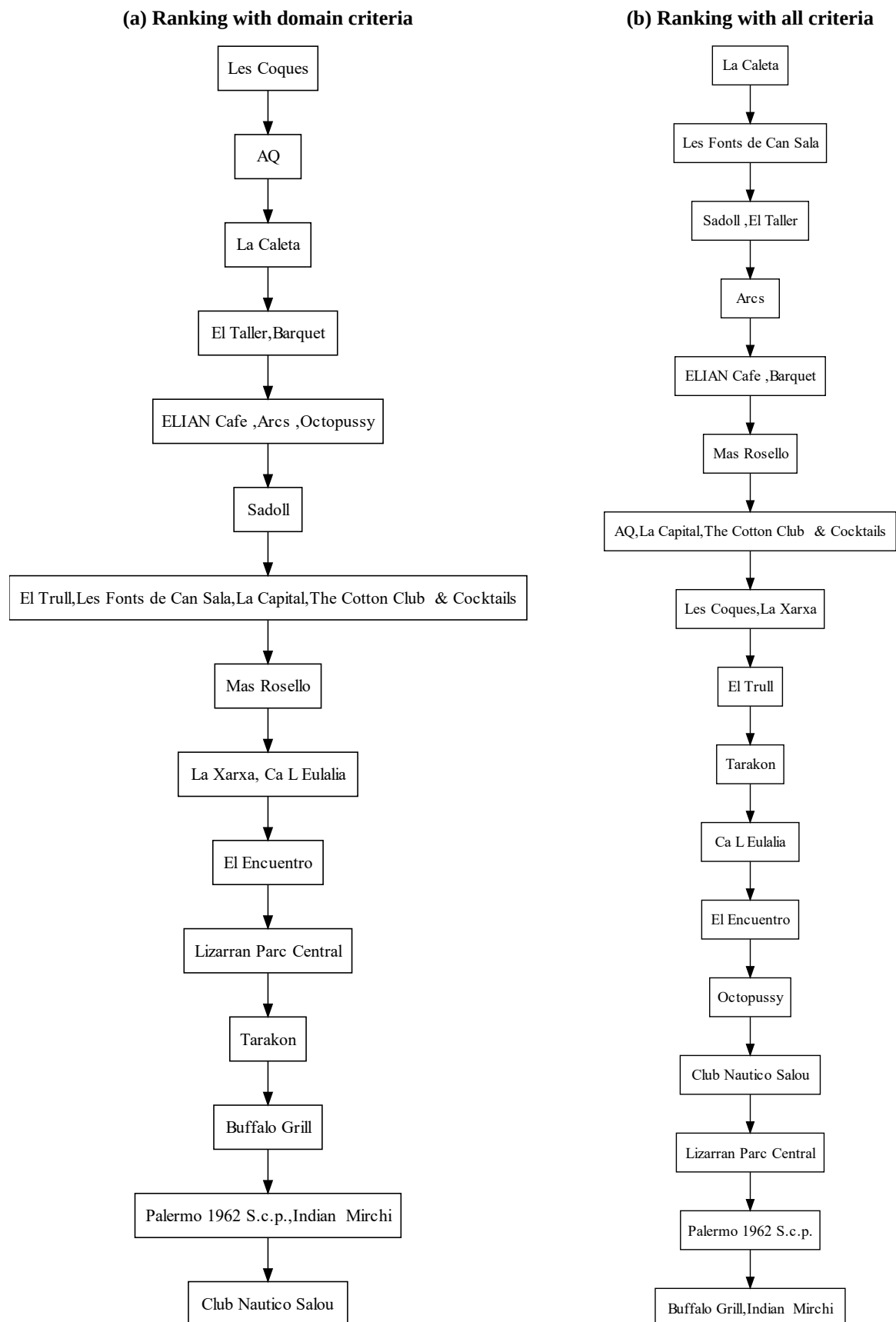


Figure 11. Profile 2: ranking result.

5. Conclusions and Future Work

Decision making is a very hard task, as it often requires the analysis of hundreds of potential alternatives defined on multiple and conflict criteria. Methods based on decision rules, utility functions or outranking methods have been proposed in the multi-criteria decision aid field. All of these methods use only the utility values of the so-called domain criteria. With the rapid growth of the Web, there is a myriad of online platforms on which users can express their opinions on the products or services offered by a company. This paper has presented SentiRank, a ranking system based on the well-known ELECTRE methodology that takes into account not only the standard domain attributes describing the alternatives but also the opinions of the users about them. A novel ABSA module has been incorporated to calculate the polarity of the users' comments about different aspects and to integrate this information as new "social" criteria. A use case with a set of restaurants has shown the important contribution of these new features in the ranking process. The degree of relevance of this contribution may be modulated with the ELECTRE weights associated with each domain and social criterion. Thus, SentiRank opens up a new line of research on the integration of the polarity of textual reviews on MCDA systems.

Concerning the future work, from a practical perspective we plan to extend the implemented system by adding more domains, such as hotels, laptops and cameras. We also plan to make the system a Web-based application and implement RESTfull APIs to allow developers and users to use the system remotely. We also plan to perform a proof-of-concept (PoC) analysis with surveys on real customers using the system. On the theoretical side, we plan to make the system multi-lingual, so that it is able to analyse reviews in other languages. We are also planning to study the possibility of using deep learning techniques in the aspect detection and aspect polarity assessment modules, as it is already being done in recommender systems [44].

Author Contributions: Concepts, M.J., A.V. and A.M.; methods, M.J.; software, M.J. and N.M.; validation, M.J., A.V. and A.M.; formal analysis, M.J., A.V. and A.M.; research, M.J.; resources and data curation, M.J. and N.M.; writing—original draft preparation, M.J.; writing—review and editing, A.M.; visualisation, M.J. and N.M.; supervision, A.M. and A.V.; funding acquisition, A.M. and A.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge the research support of URV (2018PFR-URV-B2 and Martí i Franqués PhD grant).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Floyd, K.; Freling, R.; Alhoqail, S.; Cho, H.Y.; Freling, T. How online product reviews affect retail sales: A meta-analysis. *J. Retail.* **2014**, *90*, 217–232. [\[CrossRef\]](#)
2. Simonson, I. Imperfect progress: An objective quality assessment of the role of user reviews in consumer decision making, a commentary on de Langhe, Fernbach, and Lichtenstein. *J. Consum. Res.* **2016**, *42*, 840–845. [\[CrossRef\]](#)
3. Webb, E.C.; Simonson, I. Using Reviews to Determine Preferences: How Variance in Customer-Generated Reviews Affects Choice. *ACR N. Am. Adv.* **2017**, *302*–305.
4. Shengli, L.; Fan, L. The interaction effects of online reviews and free samples on consumers' downloads: An empirical analysis. *Inf. Process. Manag.* **2019**, *56*, 102071. [\[CrossRef\]](#)
5. Saleh, K. The Importance of Online Customer Reviews [Infographic]. 2017. Available online: <https://www.invespcro.com/blog/the-importance-of-online-customer-reviews-infographic/> (accessed on 21 December 2020).
6. Ishizaka, A.; Nemery, P. *Multi-Criteria Decision Analysis: Methods and Software*; John Wiley & Sons: New York, NY, USA, 2013.
7. Bouyssou, D.; Marchant, T.; Pirlot, M.; Tsoukias, A.; Vincke, P. *Evaluation and Decision Models with Multiple Criteria: Stepping Stones for the Analyst*; Springer Science & Business Media: Berlin, Germany, 2006; Volume 86.
8. Liu, B. Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* **2012**, *5*, 1–167. [\[CrossRef\]](#)
9. Jabreel, M.; Moreno, A. SentiRich: Sentiment Analysis of Tweets Based on a Rich Set of Features. In *Frontiers in Artificial Intelligence and Applications*; Àngela Nebot, Àngela Nebot, Xavier Binefa, Ramon López de Mántaras, R.L.d.M., Eds.; IOS Press: Amsterdam, The Netherlands, 2016; Volume 288. pp. 137–146.

10. Jabreel, M.; Moreno, A.; Huertas, A. Do Local Residents and Visitors Express the Same Sentiments on Destinations Through Social Media? In *Information and Communication Technologies in Tourism 2017*; Schegg, R., Stangl, B., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 655–668.
11. Jabreel, M.; Hassan, F.; Moreno, A. Target-dependent sentiment analysis of tweets using bidirectional gated recurrent neural networks. In *Advances in Hybridization of Intelligent Methods*; Springer: Berlin, Germany, 2018; pp. 39–55.
12. Mohammad, S.; Bravo-Marquez, F.; Salameh, M.; Kiritchenko, S. Semeval-2018 task 1: Affect in tweets. In Proceedings of the 12th International Workshop on Semantic Evaluation, New Orleans, LA, USA, 5–6 June 2018; pp. 1–17.
13. Tong, S.; Koller, D. Support vector machine active learning with applications to text classification. *J. Mach. Learn. Res.* **2001**, *2*, 45–66.
14. Zhang, W.; Yoshida, T.; Tang, X. Text classification based on multi-word with support vector machine. *Knowl. Based Syst.* **2008**, *21*, 879–886. [\[CrossRef\]](#)
15. Kim, S.B.; Han, K.S.; Rim, H.C.; Myaeng, S.H. Some effective techniques for naive bayes text classification. *IEEE Trans. Knowl. Data Eng.* **2006**, *18*, 1457–1466.
16. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; ACM: New York, NY, USA, 2004; pp. 168–177.
17. Popescu, A.M.; Etzioni, O. Extracting product features and opinions from reviews. In *Natural Language Processing and Text Mining*; Springer: Berlin, Germany, 2007; pp. 9–28.
18. Scaffidi, C.; Bierhoff, K.; Chang, E.; Felker, M.; Ng, H.; Jin, C. Red Opal: Product-feature scoring from reviews. In Proceedings of the 8th ACM Conference on Electronic Commerce, San Diego, CA, USA, 11–15 June 2007; ACM: New York, NY, USA, 2007; pp. 182–191.
19. Liu, Q.; Gao, Z.; Liu, B.; Zhang, Y. Automated rule selection for aspect extraction in opinion mining. In Proceedings of the 24th International Conference on Artificial Intelligence, Buenos Aires, Argentina, 25–31 July 2015; pp. 1291–1297.
20. Jin, W.; Ho, H.H. A novel lexicalized HMM-based learning framework for web opinion mining. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; ACM: New York, NY, USA, 2009; pp. 465–472.
21. Jakob, N.; Gurevych, I. Extracting opinion targets in a single-and cross-domain setting with conditional random fields. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, MIT Stata Center, MA, USA, 9–11 October 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 1035–1045.
22. Li, F.; Han, C.; Huang, M.; Zhu, X.; Xia, Y.J.; Zhang, S.; Yu, H. Structure-aware review mining and summarization. In Proceedings of the 23rd International Conference on Computational Linguistics, Beijing, China, 23–27 August 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 653–661.
23. Choi, Y.; Cardie, C. Hierarchical sequential learning for extracting opinions and their attributes. In Proceedings of the ACL 2010, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; pp. 269–274.
24. Kiritchenko, S.; Zhu, X.; Cherry, C.; Mohammad, S. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 437–442.
25. Toh, Z.; Su, J. NLANGP: Supervised Machine Learning System for Aspect Category Classification and Opinion Target Extraction. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Denver, CO, USA, 4–5 June 2015; Association for Computational Linguistics: Denver, CO, USA, 2015.
26. Toh, Z.; Su, J. NLANGP at SemEval-2016 Task 5: Improving Aspect Based Sentiment Analysis using Neural Network Features. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016); Association for Computational Linguistics: San Diego, CA, USA, 16–17 June 2016; pp. 282–288.
27. Saías, J. Sentiue: Target and Aspect based Sentiment Analysis in SemEval-2015 Task 12. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Denver, CO, USA, 4–5 June 2015; pp. 767–771.
28. Zhang, F.; Zhang, Z.; Lan, M. Ecnu: A combination method and multiple features for aspect extraction and sentiment polarity classification. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 252–258.
29. Figueira, J.; Greco, S.; Ehrigott, M. *Multiple Criteria Decision Analysis: State of the Art Surveys*; Springer Science & Business Media: Berlin, Germany, 2005; Volume 78.
30. Figueira, J.R.; Greco, S.; Roy, B.; Słowiński, R. ELECTRE methods: Main features and recent developments. In *Handbook of Multicriteria Analysis*; Springer: Berlin, Germany, 2010; pp. 51–89.
31. Zhang, K.; Narayanan, R.; Choudhary, A.N. Voice of the Customers: Mining Online Customer Reviews for Product Feature-based Ranking. *WOSN* **2010**, *10*, 11–11.
32. Peng, Y.; Kou, G.; Li, J. A fuzzy PROMETHEE approach for mining customer reviews in Chinese. *Arab. J. Scs. Eng.* **2014**, *39*, 5245–5252. [\[CrossRef\]](#)
33. Zha, Z.J.; Yu, J.; Tang, J.; Wang, M.; Chua, T.S. Product Aspect Ranking and Its Applications. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1211–1224.

34. Kang, D.; Park, Y. Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach. *Expert Syst. Appl.* **2014**, *41*, 1041–1050. [[CrossRef](#)]
35. Liu, Y.; Bi, J.W.; Fan, Z.P. Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. *Inf. Fusion* **2017**, *36*, 149–161. [[CrossRef](#)]
36. Çalı, S.; Şebnem Yılmaz Balamam. Improved decisions for marketing, supply and purchasing: Mining big data through an integration of sentiment analysis and intuitionistic fuzzy multi criteria assessment. *Comput. Ind. Eng.* **2019**, *129*, 315–332. [[CrossRef](#)]
37. Zhang, M.L.; Li, Y.K.; Liu, X.Y.; Geng, X. Binary relevance for multi-label learning: An overview. *Front. Comput. Sci.* **2018**, *12*, 191–202. [[CrossRef](#)]
38. Tsoumakas, G.; Katakis, I. Multi-label classification: An overview. *Int. J. Data Warehous. Min. (IJDWM)* **2007**, *3*, 1–13. [[CrossRef](#)]
39. Read, J. Scalable Multi-Label Classification. Ph.D. Thesis, University of Waikato, Hamilton, New Zealand, 2010.
40. Pontiki, M.; Papageorgiou, H.; Galanis, D.; Androutsopoulos, I.; Pavlopoulos, J.; Manandhar, S. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In Proceedings of the Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; p. 27.
41. Kiritchenko, S.; Zhu, X.; Mohammad, S.M. Sentiment analysis of short informal texts. *J. Artif. Intell. Res.* **2014**, *50*, 723–762. [[CrossRef](#)]
42. Castellucci, G.; Filice, S.; Croce, D.; Basili, R. Unitor: Aspect based sentiment analysis with structured learning. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 761–767.
43. Brun, C.; Popa, D.N.; Roux, C. Xrce: Hybrid classification for aspect-based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 838–842.
44. Cheng, H.T.; Koc, L.; Harmsen, J.; Shaked, T.; Chandra, T.; Aradhye, H.; Anderson, G.; Corrado, G.; Chai, W.; Ispir, M.; et al. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, Boston, MA, USA, 15 September 2016; ACM: New York, NY, USA, 2016; pp. 7–10.