

Cross-ratio uninorms as an effective aggregation mechanism in Sentiment Analysis

Orestes Appel^{a,b}, Francisco Chiclana^a, Jenny Carter^a, Hamido Fujita^c

^aCentre for Computational Intelligence, De Montfort University, Leicester, UK

^bBissett School of Business, Mount Royal University, Calgary, Canada

^cIwate Prefectural University, Takizawa, Iwate, Japan

Abstract

There are situations in which lexicon-based methods for Sentiment Analysis (SA) are not able to generate a classification output for specific instances of a given dataset. Most often, the reason for this situation is the absence of specific terms in the sentiment lexicon required in the classification effort. In such cases, there were only two possible paths to follow: (1) add terms to the lexicon (off-line process) by human intervention to guarantee no noise is introduced into the lexicon, which prevents the classification system to provide an immediate answer; or (2) use the services of a word-frequency dictionary (on-line process), which is computationally costly to build. This paper investigates an alternative approach to compensate for the lack of ability of a lexicon-based method to produce a classification output. The method is based on the combination of the classification outputs of non lexicon-based tools. Specifically, firstly the outcome values of applying two or more non-lexicon classification methods are obtained. Secondly, these non-lexicon outcomes are fused using a uninorm based approach, which has been proved to have desirable compensation properties as required in the SA context, to generate the classification output the lexicon based approach is unable to achieve. Experimental results based on the execution of two well-known supervised machine learning algorithms, namely Naïve Bayes and Maximum Entropy, and the application of a cross-ratio uninorm operator are presented. Performance indices associated to options (1) and (2) above are compared against the results obtained using the proposed approach for two different datasets. Additionally, the performance of the proposed cross-ratio uninorm operator based approach is also compared when the aggregation operator used is the arithmetic mean instead. It is shown that the combination of non lexicon-based classification methods with specific uninorm operators improves the classification performance of lexicon-based methods, and it enables the offering of an alternative solution to the SA classification problem when needed. The proposed aggregation method could be used as well as a replacement of ensemble averaging techniques commonly applied when combining the results of several machine learning classifiers' outputs.

Keywords: Cross-ratio Uninorms, Semantic Orientation Aggregation, Hybrid Sentiment Analysis, Supervised Machine Learning, Naïve Bayes, Maximum Entropy

1. Introduction

Nowadays, understanding the opinions being conveyed by a variety of sources is an important task. The rates of volume and speed at which these waves of information are delivered to us

Email addresses: orestes.appel@my365.dmu.ac.uk (Orestes Appel), oappel@mtroyal.ca (Orestes Appel), chiclana@dmu.ac.uk (Francisco Chiclana), jennyc@dmu.ac.uk (Jenny Carter), issam@iwate-pu.ac.jp (Hamido Fujita)

are constantly growing and therefore the need for automated processing and extraction of sentiments and opinions becomes indispensable.

The most commonly used methodology to address the Sentiment Analysis (SA) problem is Supervised Machine Learning, like Naïve Bayes (NB) and Maximum Entropy (ME) techniques. In some situations more than one machine learning method can be utilised in a classification effort and a good mechanism is required to produce a final classification value that is representative of the estimation outcome of every method once they are executed independently. How to aggregate the outputs produced is not a trivial question to answer. An initial and simple approach would be to attempt to obtain an arithmetic mean as it provides a basic measure of central tendency. However, the disadvantages of an arithmetic mean are well known; such as its sensitivity to extreme values and the fact that the operator is effective only when all values being aggregated are equally important; but what if the values being utilised require a smarter aggregation? Consider, for instance, the case when two values are close to each other but a third one is far away from the other two. In this article we will present an aggregation mechanism that we believe will provide good results, as demonstrated below. Before we get to the discussion of the material, let us refresh some concepts that would help to provide some context related to SA.

Let us remember the definition of an opinion -or sentiment- as stated by Bing Liu [15]: “In an opinion we find the following items: Opinion targets (entities and their features/aspects), Sentiments (positive or negative), Opinion holders (persons who hold the opinions) and Time (when opinions are expressed). Opinions then can be: (a) Direct opinions, (b) Indirect opinions, or *Comparative Opinions*. A regular opinion is defined as a quintuple $(e_j, a_{jk}, so_{ijkl}, h_i, t_l)$ where e_j is a target entity, a_{jk} is an aspect/feature of the entity e_j , so_{ijkl} is the sentiment value of the opinion from the opinion holder h_i on feature a_{jk} of entity e_j at time t_l ”. Usually, the so-called *semantic orientation* (SO_{ijkl}) is discriminated between *positive*, *negative* or *neutral*. Let us think of using a number of SA classification methods and the need to combine together their outputs.

As our proposed approach in this paper was initially created to improve the results of a lexicon-based method that we introduced in [2], we will also refresh some of the fundamental ideas about it in this paragraph. Usually, lexicon-based methods are supported by a lexicon containing terms or words capable of conveying sentiments/opinions, most often terms belonging in the part-of-speech categories of adjectives, adverbs, verbs and nouns. Then, when a given document or sentence is analysed, the polarity values of the words/terms present in the text are utilised to calculate the associated semantic orientation, usually positive or negative, and in some cases, neutral. However, one of the challenges of lexicon-based methods is producing a classification outcome when the opinion lexicon utilised did not contained *at least* one sentiment-conveying word of those present in the sentence being processed. To address those cases when lexicon-based methods cannot issue a classification value, this paper proposes the introduction of an additional step using a uninorm based mechanism to combine the outputs of a number of supervised machine learning methods, which are capable of always producing a classification outcome, leading to an enhanced lexicon-based classification with aggregation method. Basically, in this paper we will discuss two main topics: (a) the presentation of a uninorm based aggregation method that could be used to bring together in a *compensating fashion* the results of several classification algorithms, and (b) an evaluation of the results obtained by applying the method discussed in (a) to the SA problem at the sentence level with the specific type of cross-ratio uninorm aggregation operator; mainly, as an improvement to the results presented in [2], where elaborated details can be consulted. A summarised survey on SA can be found in [1]; a complete review of the evolution of the SA field in [18] while an account of recent advances in SA techniques can be found in [8].

The remainder of this paper is organised as follows: Section 2 addresses the motivation for

introducing uninorm operators as potential aids to obtain, under special conditions, a more encompassing solution to the SA problem, whilst briefly addressing some fundamental aspects of aggregation. The definition and main properties of uninorm operators are discussed too. The specific type of cross-ratio uninorm operator used in our experiments is briefly discussed in Section 3. Section 4 presents the uninorm based method and the framework to show how to utilise it to provide a semantically sound aggregation of the outputs of two or more classification methods. The results obtained by applying the method to aggregate two well-known supervised machine learning algorithms (NB and ME) are presented, analysed and discussed in Section 5. Finally, Section 6 closes the paper with a discussion of the conclusions and the presentation of possible further work that could be addressed in the near future.

2. Preliminaries & Motivation

2.1. Preliminaries

The study of effective mechanisms for aggregation has been a central part of research in the fields of fuzzy systems and soft-computing [3, 6, 12, 20]. In [20], Rudas, Fodor & Pap mention that “the theory of fuzzy sets today uses a well developed parts of mathematics such as aggregation operations, a generalized theory of relations, generalized measure theory, etc.”. Fuzzy sets methods play a key role in many fields, of particular interest for us, in the areas of data fusion, decision-making and group decision-making. In the latter, the clear intention is to combine in a meaningful way the opinion of a number of individuals or methods.

A number of authors have performed in-depth explorations of the utilisation of aggregation functions. A very complete presentation of aggregation and aggregative uninorms can be found in the work of Yager and Rybalov [28] and Rudas and Fodor [19]. In [21] Rudas, Pap & Fodor show how key information fusion is in many complex areas like decision making, utility theory, fuzzy inference systems, robotics and vision. The authors cover aggregation functions and their fundamental properties, with four main classes of aggregation functions being identified. In [13], Jočić and Štajner-Papuga focus on pairs of binary aggregation operators on the unit interval that verify the distributivity law, which is important in the utility theory. More recently, Wu et al. [26, 27] present an interesting discussion on the use of aggregation methods for group decision-making in the specific context of social networks. The authors investigate a uninorm based approach to propagate trust through a network. In [29] the authors propose a new method to apply in group decision-making context with incomplete reciprocal preference relations. The method performs a multiplicative consistency analysis of the opinions of each expert, and provides an aggregation.

In [23–25], the mathematical modelling of the multiplicative transitivity property originally introduced by Tanino for reciprocal fuzzy preference relations is investigated and derived for the case of intuitionistic reciprocal preference relations. They use as a starting point Zadeh’s extension principle and the horizontal representation theorem of fuzzy sets based on the concept of alpha level set. Their findings assist the authors in the building of a novel consistency based induced ordered weighted averaging operator. According to these researchers, the aforementioned operator is capable of associating a higher contribution in the aggregated value to the more consistent information. In [22], Ureña et al. present an approach to decision-making based on intuitionistic preference relations, which provide a simple but flexible representation structure of experts’ preference on a set of alternative options, “while at the same time allowing to accommodate degrees of hesitation inherent to all decision making processes.” [22]. In addition, the authors introduce the concept of expert’s confidence which is based on the hesitancy degree of the reciprocal intuitionistic fuzzy preference relations. Then, they provide a group decision-making procedure, “based on a new aggregation operator that takes into account not only the experts’ consistency but also their confidence degree towards the opinion provided.”.

Meng and Chen address a new method to deal with group decision making with incomplete fuzzy preference information in [16] based on the application of an induced hybrid weighted aggregation operator. A particular feature of this aggregation operator is that “the group consistency is no smaller than the highest individual inconsistency, and the group consensus is no smaller than the smallest consensus between the individual fuzzy preference relations.” [16].

In a paper still in preparation but available in draft format on-line [9], Le Capitainé and Frélicot share their views on the reinforcement of uninorms and absorbing norms. They propose a n -ary extension of absorbing norms, defined with the support of generative functions, and its relationship with additive generating functions of uninorms. In addition, in this article the authors introduce what they consider to be a new aggregation operators (k -uninorms and k -absorbing norms). These operators are a generalization of usual uninorms and absorbing norms for which a set combination of inputs is introduced. Le Capitainé and Frélicot argue that their main ability is to provide reinforcement for contradictory inputs (as nullnorms and as opposed to uninorms). However, according to these researchers these operators still provides full reinforcement for agreeing inputs, as uninorms and as opposed to nullnorms.

2.2. Motivation

It is not uncommon for lexicon-based sentiment analysis methods to be compromised when processing sentences containing terms/words that are not in the lexicon. In situations like this there are some palliatives that could be applied, some of which are described in Appel et al. [2], like the use of a previously-generated word dictionary or vocabulary that could assist in finding an alternative solution. In general, the most common options available when a lexicon-based method cannot deliver an answer are:

- Addition of the missing word(s) into the lexicon, as suggested in [2]. This option is a valid one, provided that there is no noise introduced into the lexicon. In order to guarantee the latter, human intervention may be required, which would prevent the classification system to provide an immediate answer.
- Utilise a word-frequency dictionary, as mentioned before. This, however, may not always produce a good answer and additionally, it is typically expensive from the computational standpoint.
- Introduce a method that is not lexicon-dependent, like a machine learning algorithm as the one used by Poria et al. in [17] or another option like Naïve Bayes or Support Vector Machine.
- Select a proper aggregation technique that could smartly fusion the classification outputs of two or more algorithms that are not lexicon-dependent.

In this article we have chosen to propose the latter option among those presented above for reasons that will become evident as we progress with the presentation of the material in this article. It is important to notice, though, that the proposed technique has a significance on its own as a valid aggregation mechanism that could be utilised in many contexts. However, in this paper we have centred its use around improving the performance of an existing classification method, as introduced and proposed in [2]. Indeed, it is proposed in this study to aggregate the outcomes of two supervised machine learning techniques, namely Naïve Bayes (NB) and Maximum Entropy (ME), utilising a cross-ratio uninorm $U(x, y)$, in order to reuse the experimental results included in the work of Appel et al. [2] for comparative analysis.

We will continue expanding on this topic in Section 4, but first we will address the definition and characteristics of uninorms (Section 2.3) and the particular type of cross-ratio uninorm (Section 3) implemented in the here-proposed lexicon-based classification method with aggregation for Sentiment Analysis (SA).

2.3. Uninorm Operator

Aggregation operators are usually classified into one of the following three categories:

- (i) **Conjunctive operators** like the family of t-norm operators, which has the minimum operator as its largest element. These operators behave like a logical “and”
- (ii) **Disjunctive operators** like the family of t-conorm operators. These operators are the “dual” of conjunctive operators, and they behave like a logical “or”. The maximum operator is the smallest of all t-norms operators.
- (iii) **Compensative operators** are located between the minimum and the maximum operators, and consequently are neither conjunctive nor disjunctive. These type of operators are known as “averaging operator” and they are widely used in multi-criteria decision making problems. The arithmetic mean, the weighted mean and the ordered weighted averaging (OWA) operators are representative examples of this class.

It is worth mention *the family of uninorm operators* [28] as it does not belong fully to any of the three classes described above. Indeed, a uninorm operator, U , is defined as a mapping $U: [0, 1]^2 \rightarrow [0, 1]$ satisfying the properties:

1. Commutativity: $U(x, y) = U(y, x)$
2. Monotonicity: $U(x_1, y_1) \geq U(x_2, y_2)$ if $x_1 \geq x_2$ and $y_1 \geq y_2$
3. Associativity: $U(x, U(y, z)) = U(U(x, y), z)$
4. Identity element: $\exists e \in [0, 1] : \forall x \in [0, 1], U(x, e) = x$

Uninorm, t-norm and t-conorm operators share the commutativity, associativity and monotonicity properties. However, the set of uninorm operators has both the set of t-norm operators and the set of t-conorm operators as its subsets. Indeed, a uninorm operator with “ $e = 1$ ” becomes a t-norm operator; while a uninorm operator with “ $e = 0$ ” becomes a t-conorm operator. In general, a uninorm operator with identity element $e \in]0, 1[$ behaves like (i) a t-norm operator when all aggregated values are below e ; (ii) a t-conorm operator when all aggregated values are above e ; (iii) a compensative operator otherwise.

Notice that the semantic orientation discrimination between *positive*, *negative* or *neutral* is in accordance with the behaviour of uninorm operators. Thus, based on the above, we suggest that when a lexicon-based method, like the so-called HSC/HAC technique introduced in [2], is unable to derive the polarity of a sentence then an alternative approach could consist of implementing a uninorm operator to aggregate the polarity classification outputs, $\{x_1, x_2, \dots, x_n\}$, of non-lexicon dependent classification methods, $\{m_1, m_2, \dots, m_n\}$, respectively. Thus, the resulting aggregation would be defined by $U(x_1, x_2, \dots, x_n) = \Lambda$, where $\Lambda \in [0, 1]$ and U is an appropriate uninorm operator. In the following section, we present the class of cross-ratio uninorm operators implemented in this article.

3. Cross-ratio Aggregative Uninorm Operators

Neither t-norm operator nor t-conorm operators allow “low” values to be compensated by “high” values or viceversa. However, as explained above “uninorm operators may allow values separated by their identity element to be aggregated in a compensating way” [11].

Yager and Rybalov [28] provided the following representation of uninorms in terms of a strictly increasing continuous function of a single variable $\phi: [0, 1] \rightarrow [-\infty, \infty]$ (generator function):

$$U(x, y) = \phi^{-1} [\phi(x) + \phi(y)] \quad \forall x, y \in [0, 1]^2 \setminus \{(0, 1), (1, 0)\}.$$

such that $\phi(0) = -\infty$, $\phi(1) = +\infty$. Chiclana et al. in [10] proved that the and-like representable uninorm operator with $e = 0.5$ and $\phi(x) = \ln \frac{x}{1-x}$ [14], known as the cross-ratio uninorm,

$$U(x, y) = \begin{cases} 0, & (x, y) \in \{(0, 1), (1, 0)\} \\ \frac{xy}{xy + (1-x)(1-y)}, & \text{Otherwise.} \end{cases} \quad (1)$$

is the solution to the functional equation modelling the concept of cardinal consistency of reciprocal preference relations. The cross-ratio uninorm operator has also been utilised in the influential PROSPECTOR expert system [4]. Fodor [11] extended the cross-ratio uninorm with the identity element $e = 0.5$, so the identity element e can take on any value in $]0, 1[$:

$$U(x, y) = \begin{cases} 0, & (x, y) \in \{(0, 1), (1, 0)\} \\ \frac{(1-e)xy}{(1-e)xy + e(1-x)(1-y)}, & \text{Otherwise.} \end{cases} \quad (2)$$

Expression (2) presents the cross-ratio uninorm as an aggregation operator of two arguments. However, associativity property allows uninorm operators to fuse n (> 2) arguments:

$$U(x_1, x_2, \dots, x_n) = \begin{cases} 0, & \text{if } \exists i, j : (x_i, x_j) \in \{(0, 1), (1, 0)\} \\ \frac{(1-e)^{n-1} \prod_{i=1}^n x_i}{(1-e)^{n-1} \prod_{i=1}^n x_i + e^{n-1} \prod_{i=1}^n (1-x_i)}, & \text{Otherwise.} \end{cases} \quad (3)$$

Values in the interval $[-1, 1]$ could be used as well. Indeed, if we were interested in having semantic orientation values in $[-1, 1]$, then according to [19], there is the possibility of using the modified combining function $C: [-1, 1]^2 \rightarrow [-1, 1]$ proposed by van Melle [5]:

$$C(x, y) = \begin{cases} x + y(1-x) & , \text{if } \min(x, y) \geq 0 \\ x + y(1+x) & , \text{if } \max(x, y) \leq 0. \\ \frac{x+y}{(1-\min(|x|, |y|))} & , \text{Otherwise.} \end{cases} \quad (4)$$

Notice that C is not defined in the points $(-1, 1)$ and $(1, -1)$. However, as per Rudas and Fodor [19], rescaling function C to a binary operator on $[0, 1]$, it is possible to obtain a representable uninorm with identity element 0.5 and “as underlying t-norm and t-conorm the product and the probabilistic sum.” [19]. This result allows therefore to provide the following definition of C in $(-1, 1)$ and $(1, -1)$: $C(-1, 1) = C(1, -1) = -1$. In this article we will not be using equation (4), but the latter has been introduced in this article in an effort to show the generalisation in the method being proposed in this paper if semantic orientation values in $[-1, 1]$ were to be used.

4. Cross-ratio uninorm based SA lexicon-based methods

In this section we present how a uninorm aggregation process could be used to enhance or complement an existing lexicon-based classification method in the context of SA.

4.1. Computing Semantic Orientation (SR)

Let us assume that a given SA lexicon-based method is not capable of producing a classification score. In such a situation, we resort to the outputs of two or more supervised machine learning algorithms, and combine them together using a cross-ratio uninorm as an aggregative operator. This way, existing SA lexicon-based methods could utilise this technique as a complement strategy when its lexicon is not in a position to contribute to the generation of a classification outcome.

In order to present the usefulness of this proposed technique, we will put it at play in two different scenarios:

1. As a complement to an existing Sentiment Analysis classification method, as the one presented in Appel et al. [2].
2. As a tool in its own right that produces better results than other averaging algorithms.

In this section we will cover the case (1) above, whilst Section 5 will cover both, items (1) and (2) above.

The polarity/semantic scores for all outputs of NB and ME methods, as discussed in [2], belong in the following intervals:

- Negative values are mapped to $[0, 0.4999]$.
- Positive values are mapped to $[0.5001, 1]$.

Now, all scores belong in the unit interval, leaving the value of 0.5000 to represent Objective/Neutral semantic orientation. The value of 0.5000 corresponds with the identity element e introduced in equations (1) and (2) in Section 3. As a consequence, the new introduced polarity spectrum maps to two symmetrical half unit-interval ranges and their values are ready to be aggregated by a cross-ratio uninorm having identity element $e = 0.5000$.

4.2. The proposed aggregation process

Let us recall that the outputs of NB and ME are *numbers* that belong in the unit interval, and the cut off point is given by values that are either greater than or lower than 0.5.

If Output(NB) > 0.5 **Then** Semantic Orientation is *Positive*.

If Output(NB) < 0.5 **Then** Semantic Orientation is *Negative*.

If Output(ME) > 0.5 **Then** Semantic Orientation is *Positive*

If Output(ME) < 0.5 **Then** Semantic Orientation is *Negative*.

The above If-Then clauses satisfy our criteria of the identity element being $e = 0.5$. In the case that a lexicon-based classification method, as in [2], is occasionally incapable of producing a classification value as a consequence of some of the limitations of lexicon-based approaches, then the proposed enhanced alternative option could be put at play, as described below:

1. Collect the outputs of NB (ONB)
2. Collect the outputs of ME (OME)
3. For each sentence, feed both outputs -ONB and OME- to the cross-ratio uninorm presented above in equation (1), which will produce a value (*Result*) for each sentence
 - (a) **If** *Result* > 0.5 **Then** Semantic Orientation = *Positive*
 - (b) **If** *Result* < 0.5 **Then** Semantic Orientation = *Negative*
 - (c) **IIf** *Result* = 0.5 **Then** Semantic Orientation = *Neutral*

The enhanced SA lexicon-based classification with aggregation method is graphically depicted in Fig. 1, with the shaded area corresponding to the aggregation of the classification outputs of the NB and ME algorithms as proposed above. Notice that this proposed improvement *ensures* that our Hybrid Advanced Classification Method, as described in [2], will get enhanced as it will always be in a position to produce a classification output and neither off-line addition of missing word(s) into the lexicon nor a computationally expensive word-frequency dictionary are needed.

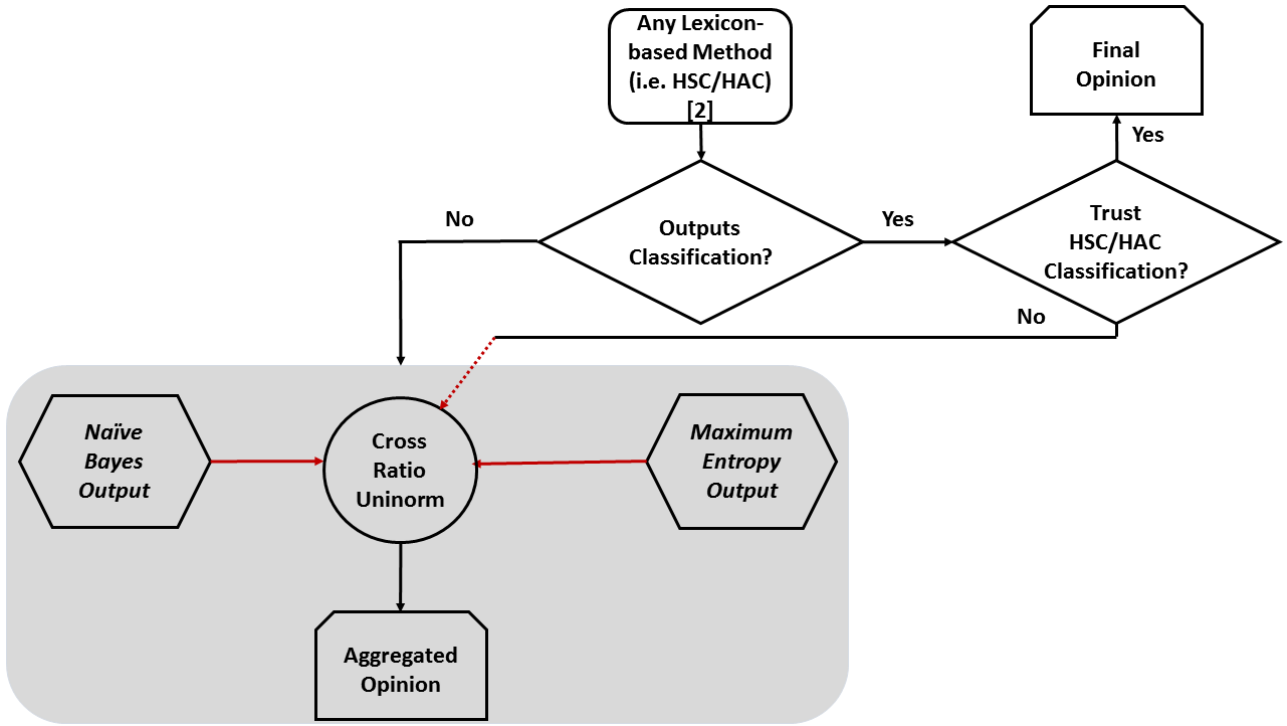


Figure 1: Cross-ratio Uninorm Aggregation

5. Experimental results

In the method proposed by Appel et al. in [2], when the sentiment lexicon did not contain the necessary terms/words to produce a classification output when processing a given sentence, there were only two possible paths to follow: (a) add terms to the lexicon (off-line process), or (b) use the services of a Word-frequency dictionary (on-line process) that is computationally costly to build. The latter method has been utilised in our HSC method [2], as a last resort, when the necessary terms for producing a classification are not present in our sentiment lexicon. Experimental results are reported in this section regarding the performance of the alternative path proposed in this paper based on the cross-ratio uninorm operator.

5.1. Datasets utilised

We make use of the Movie Review Dataset provided by Pang and Lee and available at <http://www.cs.cornell.edu/people/pabo/movie-review-data/>. In addition, we utilise a dataset containing Twitter data, *Sentiment140*, which is available at <http://help.sentiment140.com/for-students>.

5.2. Results for the application of Cross-ratio Uninorm Aggregation to test datasets

In order to assess the validity of the alternative path based on the cross-ratio uninorm proposed in this article, experiments were carried out to compare:

1. The cross-ratio uninorm performance as an aggregation tool of the classification outputs when applying the NB and ME algorithms against the performance achieved using the *arithmetic mean* aggregation instead, and the classification performance using exclusively the services of a *word-frequency dictionary* (Subsection 5.2.1);
2. The performance of the the cross-ratio uninorm implementation in the HSC lexicon-based method [2] against the performances of the HSC lexicon-based method with the off-line

addition of missing-words to the lexicon and with the Word-frequency dictionary, respectively, when the lexicon cannot respond (Subsection 5.2.2). The performance of the Cross-ratio uninorm method when embedded in our hybrid method (HSC) as a complement

5.2.1. Cross-ratio uninorm against two other possible techniques

Tables 1 and 2 show comparative results of the cross-ratio uninorm aggregation the classification outputs when applying the NB and ME algorithms against two methods: (a) the arithmetic mean, and (b) the classification outputs obtained by using the Word-frequency dictionary described in [2]. Experiments have been performed for both the Movie DB dataset (10,662 occurrences, the complete set) and the Twitter dataset (15,000 occurrences).

Alternative Method	Accuracy	Precision	Recall	F1-score
Arithmetic Mean	0.46	0.52	0.47	0.49
Word-frequency Dictionary	0.58	0.66	0.57	0.61
Cross-ratio Uninorm (NB & ME)	0.66	0.67	0.64	0.66

Table 1: Method Vs. Indicators (Movie DB: 10,662 sentences)

Alternative Method	Accuracy	Precision	Recall	F1-score
Arithmetic Mean	0.50	0.57	0.43	0.49
Word-frequency Dictionary	0.56	0.69	0.56	0.62
Cross-ratio Uninorm (NB & ME)	0.75	0.78	0.82	0.80

Table 2: Method Vs. Indicators (Twitter dataset: 15,000 sentences)

In this comparison, the cross-ratio uninorm (NB & ME) comes ahead of the other two algorithms in all four performance indicators. Particularly, the recall indicator (how many of the true positives sentences were found) displayed by the cross-ratio uninorm is the highest of all by a significant margin. In the next section, when the sentiment lexicon cannot respond, the cross-ratio uninorm (NB & ME) technique enhancement to our HSC [2] will be compared against the HSC lexicon-based method with the off-line addition of missing-words to the lexicon and with the Word-frequency dictionary, respectively.

5.2.2. Cross-ratio uninorm as an enhancer of our hybrid method

In the hybrid model we presented in [2] we did show that during the first pass of the proposed hybrid algorithm, there were sentences that could not be classified as the sentiment lexicon did not count with the required terms/words in order to produce a classification outcome. Two approaches to circumvent this problem were proposed in [2]: (a) the use of a word-frequency dictionary that served its purpose but it has a negative aspect in that its creation involves an algorithm of complexity $\mathcal{O}(n^2)$, where n is the numbers of words in the dataset being utilised; or (b) incorporate new terms into the dictionary, which could not be done interactively and required an expert human intervention.

The next step in our experimental phase is to study and analyse the performance of the cross-ratio uninorm (NB & ME) when embedded in our HSC method [2]. To force the application of the cross-ratio uninorm, terms were randomly removed from the sentiment lexicon leading to a number of sentences that could not be classified using the HSC method (1,337 sentences in total). Hence, the services of: (1) the Word-frequency Dictionary; (2) the Cross-ratio Uninorm; and (3) off-line addition of missing terms to the sentiment lexicon, were demanded for those 1,337 sentences. The obtained results are presented in Table 3 (for Movie Dataset) and Table 4 (for Twitter Dataset).

Method	Accuracy	Precision	Recall	F1-score
(1) Hybrid using Word Dictionary	0.66	0.64	0.77	0.69
(2) Hybrid using Cross-ratio Uninorm	0.74	0.71	0.82	0.76
(3) Hybrid with word-addition enabled	0.76	0.73	0.83	0.77

Table 3: All Hybrid methods derived from HSC [2] - Movie Dataset

We see that the cross-ratio uninorm approach performs second-best in the group, achieving results that are very close to those attained by HSC with word-addition enabled (only 2% below in Accuracy and Precision). In third place we get the Word-frequency Dictionary that is 8% and 7% worse than the cross-ratio uninorm for Accuracy and Precision, respectively.

We repeated the same experiment using this time the Twitter Dataset, obtaining the results presented in Table 4, which basically confirms the previous results and analysis. We can say with confidence that the alternative option when the lexicon cannot offer a solution based on the Cross-ratio Uninorm constitutes an excellent alternative at a very low cost, both computationally and people-wise.

Method	Accuracy	Precision	Recall	F1-score
(1) Hybrid using Word Dictionary	0.77	0.74	0.84	0.79
(2) Hybrid using Cross-ratio Uninorm	0.86	0.81	0.93	0.87
(3) Hybrid with word-addition enabled	0.88	0.84	0.94	0.89

Table 4: All Hybrid methods derived from HSC [2] - Twitter Dataset

6. Conclusions and further work

Cross-ratio uninorm operators can certainly play a significant role in aggregating the opinions of a number of classification systems in a more balanced way, compensating when required for specific data traits as discussed in section 2.3, behaving like a *conjunctive*, *disjunctive* or *compensatory* operator as required.

If we recall our initial motivation, for those cases when a lexicon-based SA method cannot produce a classification output (there are terms required for the analysis that are absent from the sentiment lexicon) we could as an alternative option use the services of a Word-frequency Dictionary or add to the lexicon (off-line) the missing words required to complete the analysis. However, adding words off-line is expensive because it requires the knowledge of an expert and it is time consuming, which means that the method cannot produce an answer immediately, i.e. it prevents its automation. The other alternative method involves the creation of the Word-frequency Dictionary, which it is also computationally expensive ($\mathcal{O}(n^2)$) plus time searching through it). In contrast to these approaches, the alternative cross-ratio uninorm approach is easy to implement and not computational expensive. In addition, it performs much better performer than the Word-frequency Dictionary. Although the proposed cross-ratio uninorm approach performs slightly worse than the off-line addition of the missing words to the lexicon, it has the advantage of being less time consuming and allows to automate the whole SA process. This situation provides us with options:

1. If one can afford the costs of adding missing terms to the sentiment lexicon and it is possible to wait for a more precise answer, then HSC as presented in [2] is the best choice.
2. If one is urged to provide an answer immediately, then there is the convenient alternative of using the cross-ratio uninorm approach presented in this article

It is a matter of a compromise, between off-line extra time to look for the required words and getting them in the lexicon, and providing an immediate answer with a potential slight lesser accuracy.

In order to put things in perspective, let us remember that many commercial software packages in the realm of machine learning provide the option of utilising the so-called ensemble averaging methods. Typically, this technique works by combining previously created methods in order to produce a desired output. Usually the steps are: (i) obtain the outputs of N methods, (ii) separately, train each model, and (iii) combine the method outputs and average their values. In some cases, a slightly more complex approach is followed, and the ensemble averaging is performed as $\tilde{y}(x; \alpha) = \sum_{j=1}^N \alpha_j y_j(x)$, where each method output is y_j , α values represent a set of weights, and N is the number of methods being considered. This version corresponds to a weighted sum instead of a mere average. However, as it has been shown in the experimental results section, the proposed uninorm based method performs better than standard average functions. As such, we suggest that using a uninorm approach would provide better results. In addition, if it is true that the proposed technique was introduced as a complement to the lexicon-based classification method presented in [2], we believe that the cross-ratio uninorm described in this article could be utilised as well as a more efficient and focused ensemble method where the semantic of the aggregation represents a symmetric aggregative effect.

In terms of further work, there are two avenues that could be pursued in the short-term:

- In addition to obtaining the aggregation already mentioned among the classification methods, we could incorporate the level of trust that one has on each method $\{M_1, M_2, \dots, M_n\}$, ensuring that those better established and proven methods carry more weight (as we would do when pondering the opinions of a number of people, depending on how much we *trust* each of them).
- There are multiple uninorm operators available (which stem from the research field of *decision making theory*) and some of them are highly flexible depending on the selection of the appropriate quantifier or function. As such, we would like to explore further additional options that could potentially provide even better results. Especially, around the possibility of utilising equation 4 in aggregating words polarity scores and sub-sentences polarity scores in an approach to produce a semantic orientation score without the need for using polarity labels (Positive/Negative) and using the polarity interval $[-1, 1]$ instead of $[0, 1]$. Of course, such changes would require the introduction of modifications to the existing sentiment lexicons.

References

- [1] Orestes Appel, Francisco Chiclana, and Jenny Carter. Main Concepts, State of the Art and Future research Questions in Sentiment Analysis. *Acta Polytechnica Hungarica - Journal of Applied Sciences*, 12(3):87–108, 2015.
- [2] Orestes Appel, Francisco Chiclana, Jenny Carter, and Hamido Fujita. A hybrid approach to the sentiment analysis problem at the sentence level. *Knowledge-Based Systems*, 108:110 – 124, 2016.
- [3] Bernadette Bouchon-Meunier (Ed.) *Aggregation and Fusion of Imperfect Information*, Springer-Verlag Berlin Heidelberg, 1998.
- [4] Bernard De Baets and János Fodor. Van Melle’s combining function in MYCIN is a representable uninorm: An alternative proof. *Fuzzy Sets Systems*, 104:133–136, 1999.

- [5] Bruce G. Buchanan and Editors Edward H. Shortliffe (Ed.) *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, Reading, MA, 1984.
- [6] Francisco Javier Cabrerizo, Francisco Chiclana, Rami Al-Hmouz, Ali Morfeq, Abdullah Saeed Balamash and Enrique Herrera-Viedma. Fuzzy decision making and consensus: challenges. *Journal of Intelligent & Fuzzy Systems*, 29(3), 1109–1118, 2015.
- [7] Erik Cambria, Daniel Olsher, and Dheeraj Rajagopal. Senticnet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, 1515–1521, 2014.
- [8] Erik Cambria, Bjorn Schuller, Yunqing Xia, and Catherine Havasi. New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, 28(2):15–21, March 2013.
- [9] Hoel Le Capitainé and Carl Frélicot. On the reinforcement of uninorms and absorbing norms. Working paper. Published in HAL (an open archive where authors can deposit scholarly documents from all academic fields, reachable at <https://hal.archives-ouvertes.fr>), Aug 2016.
- [10] Francisco Chiclana, Enrique Herrera-Viedma, Sergio Alonso, and Francisco Herrera. Cardinal consistency of reciprocal preference relations: A characterization of multiplicative transitivity. *IEEE Transactions on Fuzzy Systems*, 17(1):14–23, 2009.
- [11] János Fodor. On Rational Uninorms. In *Proceedings of the First Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence, Herlany, Slovakia, February 12-14, 2003*, 139–147, 2003.
- [12] János Fodor. Aggregation Functions in Fuzzy Systems. In *Fodor, János and Kacprzyk, Janusz (Ed.). Aspects of Soft Computing, Intelligent Robotics and Control, Springer Berlin Heidelberg*, 25–50, 2009.
- [13] Dragan Jočić and Ivana Štajner-Papuga. Distributivity equations and Mayor’s aggregation operators. *Knowledge-Based Systems*, 52:194–200, Nov 2013.
- [14] Erich Peter Klement, Radko Mesiar, and Endre Pap. On the relationship of associative compensatory operators to triangular norms and conorms. *Int. J. Uncertainty, Fuzziness Knowledge-Based Systems*, 4:129–144, 1996.
- [15] Bing Liu. *Tutorial: Sentiment Analysis Tutorial Given at AAAI-2011, Monday, August 8 2011, San Francisco, USA. Twenty-Fifth Conference on Artificial Intelligence (AAAI-11), held in San Francisco, California at the Hyatt Regency San Francisco*, August 7–11 2011.
- [16] Fanyong Meng and Xiaohong Chen. A new method for group decision making with incomplete fuzzy preference relations. *Knowledge-Based Systems*, 73:111–123, Jan 2015.
- [17] Soujanya Poria, Erik Cambria, Grégoire Winterstein, and Guang-Bin Huang. Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, 69:45–63, 2014.
- [18] Kumar Ravi Ravi and Vadlamani Ravi. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 89(C):14–46, Nov 2015.

- [19] Imre J. Rudas and János Fodor. Information Aggregation in Intelligent Systems Using Generalized Operators. *International Journal of Computers, Communications & Control*, I(1):47–57, 2006.
- [20] Imre J. Rudas, Endre Pap, and János Fodor. Editorial: Special issue on Advances in fuzzy knowledge systems: Theory and application. *Knowledge-Based Systems*, 38:1–2, Jan 2013.
- [21] Imre J. Rudas, Endre Pap, and János Fodor. Information aggregation in intelligent systems: An application oriented approach. *Knowledge-Based Systems*, 38:3–13, Jan 2013.
- [22] Raquel Ureña, Francisco Chiclana, Hamido Fujita, and Enrique Herrera-Viedma. Confidence-consistency driven group decision making approach with incomplete reciprocal intuitionistic preference relations. *Knowledge-Based Systems*, 89:86–96, Nov 2015.
- [23] Jian Wu and Francisco Chiclana. Multiplicative consistency of intuitionistic reciprocal preference relations and its application to missing values estimation and consensus building. *Knowledge-Based Systems*, 71:187–200, Nov 2014.
- [24] Jian Wu, and Francisco Chiclana. Visual information recommendation simulation and attitudinal prioritisation method for group decision making with triangular fuzzy complementary preference relations. *Information Sciences* 279, 716–736, (2014).
- [25] Jian Wu, Francisco Chiclana, and Huchang Liao. Isomorphic multiplicative transitivity for intuitionistic and interval-valued fuzzy preference relations and its application in deriving their priority vectors. *IEEE Transactions on Fuzzy Systems*, 2016. doi: 10.1109/TFUZZ.2016.2646749.
- [26] Jian Wu, Francisco Chiclana, and Enrique Herrera-Viedma. Trust based consensus model for social network in an incomplete linguistic information context. *Applied Soft Computing* 35: 827–839, 2015.
- [27] Jian Wu, Ruoyun Xiong, and Francisco Chiclana. Uninorm trust propagation and aggregation methods for group decision making in social network with four tuple information. *Knowledge-Based Systems*, 96:29–39, 2016.
- [28] Ronald R. Yager and Alexander Rybalov. Uninorm aggregation operators. *Fuzzy Sets and Systems*, 80(1):111 – 120, 1996.
- [29] Xiaoxiong Zhang, Bingfeng Ge, Jiang Jiang, and Yuejin Tan. Consensus building in group decision making base on multiplicative consistency with incomplete reciprocal preference relations. *Knowledge-Based Systems*, 106:96–104, Aug 2016.