

# A Hybrid Approach to Sentiment Analysis

Orestes Appel, Francisco Chiclana and Jenny Carter  
Centre for Computational Intelligence (CCI)  
De Montfort University  
Leicester, UK

Email: orestes.appel@email.dmu.ac.uk, {chiclana, jennyc}@dmu.ac.uk

Hamido Fujita  
Iwate Prefectural University (IPU)  
Takizawa, Iwate, Japan  
Email: issam@iwate-pu.ac.jp

**Abstract**—This contribution presents a hybrid approach to Sentiment Analysis (SA) encompassing the use of semantic rules, fuzzy sets, unsupervised machine learning techniques and a sentiment lexicon improved with the support of Senti-WordNet. A Hybrid Standard Classification is first carried out, which is further enhanced into a Hybrid Advanced approach incorporating linguistic classification of semantic polarity modelled using fuzzy sets. The mechanism of the new SA methodology is illustrated by applying it to compute the polarity of a given sentence and to a benchmarking publicly available dataset: the Movie Review Dataset.

## I. INTRODUCTION

Understanding the emotions being conveyed by a given source, may it be a tweet, a document, a report, a blog, a segment of a politician speech, etc., has proved to be an important activity for humans. However, when volumes of *opinions* are many, human processing becomes a challenge, hence, the need for automated processes to extract sentiments from a variety of sources that keep growing in volume, complexity and diversity. Michelle de Haaff provided the following definition of Sentiment Analysis (SA) in her CustomerThink (2010) *blog* entry entitled, Sentiment Analysis, Hard But Worth It: "...classifying the polarity of a given text at the document, sentence, or feature or aspect level, whether the expressed opinion in a document, a sentence or an entity feature or aspect is positive, negative, or neutral.". Usually, approaches to address the SA problem, which is often called Opinion Mining problem, belong either to the category of Supervised or Unsupervised Machine Learning.

However, it seems that fuzzy sets, considering their mathematical properties and their ability to deal with vagueness and uncertainty, are well-equipped as well to model sentiment-related problems. As we know, fuzzy relations have been extensively used in disciplines as dissimilar as linguistic [1], clustering and decision-making, among many others. Thus, a combination of techniques could be more effective at succeeding at addressing the Sentiment Analysis challenges. In the next paragraph we will address our motivation to explore this realm of possibilities.

Dzogang et al. stated in [2] that usually authors refer mainly to psychological models when attacking the sentiment analysis problem. However, other models may

be successful as well in this domain. As per Dzogang et al. "...it must be underlined that some appraisal based approaches make use of graduality through fuzzy inference and fuzzy aggregation for processing affective mechanisms ambiguity and imprecision ...". When dealing with Sentiment Analysis, Bing Liu [3], one of the main world experts in this area, mentions that "...We probably relied too much on Machine Learning.". Hence, the following arguments combined together sparked:

- The concept of graduality for polarity intensity expressed through fuzzy sets
- The idea that other alternatives -like a number of Natural Language Processing techniques-, besides Supervised Machine Learning, may be viable as well when extracting sentiment from text
- The positive contribution that semantic rules and a solid opinion lexicon can have in identifying subjectivity in SA

For a complete review of the evolution of the Sentiment Analysis field, please refer to the work of Appel et al. [4].

## II. RESEARCH METHODOLOGY

The research methodology to be used is discussed from two different perspectives: the process to follow and the data to use for measuring the performance of the proposed SA solution.

### A. The Process

In order to measure success, the proposed solution should perform *same or better* than today's most accepted and commonly used solutions. In the specific case of the SA problem, the proposed solution will be compared in the future against two Supervised Learning methods that enjoy a high level of acceptance and credibility in the classification research community and that are relatively straightforward to implement: Naïve Bayes (NB) and Maximum Entropy (ME). We will focus on two standard aspects at the *sentence level*:

- 1) *Subjectivity determination*: been able to tell an opinion from a fact (*Qualification*)
- 2) *Opinion polarity (graduality)*: assigning a value to an opinion inside a given range [positive, negative, neutral, etc.] (*Quantification*)

In this paper we will focus only on the architecture of the SA Hybrid Methodology and its application to a public dataset, whilst showing some initial results. A full comparison with other methods will be reported in the near future.

### B. The Data

A natural question to answer at this point is what data to use to benchmark our results. As there are many studies on the subject, it was decided to examine the existing cases-databases publicly available and the results obtained with those data sets coming from the most recognised research figures in the SA world. The dataset chosen for our study is owed to Pang and Lee. They had published in [5] datasets that were utilised in SA experiments, which results were addressed and discussed in [6, 7, 8]. There are two main reasons why these datasets are appropriate for this study: (i) they are named after version numbers that can be cited and that have been used in previous experiments that show results publicly, and (ii) the datasets are sub-divided into categories, namely, sentiment polarity datasets, sentiment scale datasets and subjectivity datasets. As such, it seems adequate to use the Movie Review Dataset provided by Pang & Lee and available at <http://www.cs.cornell.edu/people/pabo/movie-review-data/>. The dataset used in our experiments corresponds to v1.0 of the sentence polarity datasets, `rt-polaritydata.tar.gz`, which was used in the experiments described in [6]. The dataset contains 5,331 positive snippets and 5,331 negative ones.

### C. Most commonly used performance indicators in SA

It has become customary to evaluate the performance of sentiment classification systems utilising the following four indexes, as defined in [9]:

- *Accuracy*: the portion of all true predicted instances against all predicted instances
- *Precision*: the portion of true positive predicted instances against all positive predicted instances
- *Recall*: the portion of true positive predicted instances against all actual positive instances
- *F1-score*: a harmonic average of precision and recall

For more details on these performance indicators, please look at Sadegh and Othman [9].

The rest of this article will present our proposed Hybrid Method for Sentiment Analysis covering its three main components, namely: (i) the Sentiment/Opinion Lexicon, (ii) Negation handling and Semantic Rules and (iii) the use of Fuzzy Sets for increasing granularity in the polarity determination. In closing, we will explain the mechanics behind the Hybrid Process and will share our conclusions.

## III. A PROPOSED HYBRID APPROACH TO THE SA PROBLEM AT THE SENTENCE LEVEL

Let us discuss a little further what exactly we do mean by utilising a *Hybrid Approach* that is key to our proposed solution. Our intention is to manage *hybrid concepts* at two different levels:

- 1) The methods employed by the sentiment classifier itself
- 2) The techniques used to build key components in our hybrid approach, like the creation and population of the Sentiment/Opinion Lexicon, the building of a dictionary of words including previous occurrences of sentiment/opinion cases and associated occurrence frequencies and the incorporation of other techniques (i.e. fuzzy sets)

The following paragraphs will discuss in detail the aforementioned aspects and will cover the different components of our proposed hybrid solution.

### A. Component 1: The Sentiment/Opinion Lexicon

Prof. Liu [10] compiled an Opinion Lexicon that "...does include a list of positive and negative opinion words or sentiment words for English (around 6,800 words) compiled over many years ...". The compilation time span goes back to their initial work in the field. This Opinion Lexicon will be used as a starting point and it will be enriched in a number of ways, including a new structure and organisation more adequate to the research approach here proposed. Part of the reasoning behind using Liu's Lexicon is to re-use data resulting from a good quality massive effort in words compilation that already took place and as a point of *commonality* with previous research efforts for benchmarking purposes. In generating our own Opinion Lexicon:

- 1) We have utilised the *opinion-conveying-words* that are part of the Opinion Lexicon used by Hu and Liu in [10]. They correspond to two lists, one containing *positive-meaning words* and the other one containing *negative-meaning words*. They include *only* four elements of Part-of-Speech (PoS) that have been proved to be capable of delivering opinions [11, 12, 13, 14]: nouns, verbs, adjectives and adverbs.
- 2) We have used SentiWordNet [15, 16], by assigning a *polarity or valence score* to words carrying opinion sense.
- 3) We have combined both elements in (1) and (2) above. As such, we have substituted the words in Liu's original Opinion Lexicon for their Synset-equivalent in SentiWordNet (at least partially as explained in (4) below). By doing this, we have added a *positive score* (PSC) and a *negative score* (NSC) to the existing words in Liu's Opinion Lexicon. The *objectivity score* (COBJ) is calculated as:  $COBJ = 1 - (PSC + NSC)$ . The PSCs and the NSCs

for the words in the lexicon are *taken* from SentiWordNet. We must keep in mind that  $0 \leq PSCs \leq 1$  and  $0 \leq NSCs \leq 1$ .

- 4) Not every word in Liu’s Opinion Lexicon is present in SentiWordNet. Hence, for those absent words we have chosen to keep them in the new Opinion Lexicon, but they do not have polarity scores associated neither the proper PoS tag.

It is important to notice that WordNet has been collected, prepared and documented by human being experts in the subject at hand, whilst SentiWordNet has followed a different approach that involves machine learning algorithms and machine-assisted analysis, although some manual work has been performed as well. In a way, as we use SentiWordNet extensively, we are inheriting the supervised machine learning techniques utilised to compile SentiWordNet. Here is the description of the attributes of our proposed lexicon:

Word: word in the lexicon (entries).

PoS: part of speech (n=noun; v=verb; a=adjective; r=adverb; s=adjective satellite).

PSC: Positive Score as taken from SentiWordNet [15].

NSC: Negative Score as taken from SentiWordNet [15].

COBJ: Calculated Objectivity Score.

VDX: Versioning index for identifying/managing synonyms (future use).

UPDC: Update Counter to keep track of every time a given entry in the lexicon is updated.

If there is *missing* information, a label indicating so is placed in the proper place in the sentiment lexicon for posterior use. As mentioned in the previous section, we have as a base-case a subset of opinion-conveying words for which we could not find a synset in SentiWordNet. For those cases we have initially assumed a convention, so those elements can be easily identified later on when needed. This is part of the process of *increasing the knowledge of the Opinion Lexicon*, which is out of the scope of this article.

### B. Component 2: Semantic Rules (SR), Negation Handling and a Lexicon Approach to SA

A number of authors, among them [17, 18, 19], have pointed out to the fact that negation and the use of specific part-of-speech particles, like *but*, *despite*, *unless*, and *others*, and composition of functions -i.e.  $compose(v1, v2)$ - could affect the final outcome of a classification exercise.

In this section we will address those cases for which we must define certain rules to model the problem of SA in a more accurate fashion. Let us look at this in more detail. Firstly we show the rules and then we discuss the main aspects around Table I and Table II. It is important to notice that this approach or methodology has been evolving in time, as researchers keep improving

the rules and elements considered to be important in the determination of Semantic Orientation (SO). The rules here presented are very well organised, easy to read and certainly more complete than others we have had the chance to review. Regardless, some rules strategies were needed to be put in place as the order of the different part-of-speech play a key role in the semantic of a given sentence. Despite the apparent completeness of the Semantic Rules shown and described by Xie et al.[19], we have incorporated *two new rules* for managing particular part-of-speech particles: the particle **while** and the particle **however**, which were not included in the original set of rules provided in [19].

TABLE I  
SEMANTIC RULES FOR OUR HYBRID SYSTEM

Rule No.	Semantic Rule
Rule 1	Polarity (not $var_k$ ) = -Polarity ( $var_k$ )
Rule 2	Polarity ( $NP_1$ of $NP_2$ ) = $Compose(NP_1, NP_2)$
Rule 3	Polarity ( $NP_1 VP_1$ ) = $Compose(NP_1, VP_1)$
Rule 4	Polarity ( $NP_1$ be $ADJ$ ) = $Compose(ADJ, NP_1)$
Rule 5	Polarity ( $NP_1$ of $VP_1$ ) = $Compose(NP_1, VP_1)$
Rule 6	Polarity ( $ADJ$ to $VP_1$ ) = $Compose(ADJ, VP_1)$
Rule 7	Polarity ( $VP_1 NP_1$ ) = $Compose(VP_1, NP_1)$
Rule 8	Polarity ( $VP_1$ to $VP_2$ ) = $Compose(VP_1, VP_2)$
Rule 9	Polarity ( $ADJ$ as $NP$ ) = $1_{(Polarity(NP=0))} \cdot Polarity(ADJ) + 1_{(Polarity(NP \neq 0))} \cdot Polarity(NP)$
Rule 10	Polarity (not as $ADJ$ as $NP$ ) = -Polarity ( $ADJ$ )
Rule 11	If sentence contains <b>but</b> , disregard all previous and only take the sentiment of the part after <b>but</b>
Rule 12	If sentence contains <b>despite</b> , only take the sentiment of the part before <b>despite</b>
Rule 13	If sentence contains <b>unless</b> , and <b>unless</b> is followed by a negative clause, disregard the <b>unless</b> clause
Rule 14 - (New)	If sentence contains <b>while</b> , disregard the sentence following the <b>while</b> and take the sentiment only of the sentence that follows the one after the <b>while</b>
Rule 15 - (New)	If sentence contains <b>however</b> , disregard the sentence preceding the <b>however</b> and take the sentiment of the sentence that follows the one after the <b>however</b>

1) *Negation Handling*: According to the well-known researcher Christopher Potts, Stanford University, Linguistics Department, “Sentiment words behave very differently when under the semantic scope of negation” [20]. Dr. Potts notices that the so-called *Weak* (mild) words such as *good* and *bad* behave like their opposites when negated (*bad* = not *good*, *good* = not *bad*), whilst “*Strong* (intense) words like *superb* and *terrible* have

TABLE II  
COMPOSE FUNCTIONS REFERENCED IN TABLE I

Compose Functions	Algorithms
Compose1 (arg1, arg2)	1. Return $-Polarity(arg2)$ if $arg1$ is negation.
	2. Return $Polarity(arg1)$ if $(Polarity(arg1) = Polarity(arg2))$ .
	3. Otherwise, return the majority term polarity in $arg1$ and $arg2$ .
Compose2 (arg1, arg2)	1. Return $Polarity(arg2)$ if $arg1$ is negative and $arg2$ is not neutral.
	2. Return $-1$ if $arg1$ is negative and $arg2$ is neutral.
	3. Return $Polarity(arg2)$ if $arg1$ is positive and $arg2$ is not neutral.
	4. Return $2 * Polarity(arg1)$ if $Polarity(arg1) = Polarity(arg2)$ .
	5. Return $Polarity(arg1) + Polarity(arg2)$ if $arg1$ is positive and $arg2$ is neutral.
	6. Return $Polarity(arg1) + Polarity(arg2)$ if $arg2$ is positive and $arg1$ is neutral.
	7. Otherwise, return $0$ .

very general meanings under negation.”. According to Potts: “...not superb is consistent with everything from horrible to just-shy-of-superb, and different lexical items favor different senses. These observations suggest that it would be difficult to have a general *a priori* rule for how to handle negation. It doesn’t just turn good to bad and bad to good. Its effects depend on the words being negated. An additional challenge for negation is that its expression is lexically diverse and its influences are far-reaching (syntactically speaking)”. As a consequence, a method addressing the behaviour just described is the one favoured by Dr. Potts. This technique is due to Das and Chen [21] and Pang, Lee, and Vaithyanathan [8]. The method works as follows (notice it uses regular expressions):

- 1) *Negation marking*: Append a `_NEG` suffix to every word appearing between a negation and a clause-level punctuation mark.
- 2) *Negation*: A negation is any word matching the following regular expression  
(?:  
 $\wedge$  (?:(never | no | nothing | nowhere | noone | none | not |  
havent | hasnt | hadnt | cant | couldnt | shouldnt

|  
wont | wouldnt | dont | doesnt | didnt | isnt |  
arent | aint  
)\$

)  
|  
n’t

- 3) *Clause-level punctuation*: A clause-level punctuation mark is any word matching the following regular expression:  $\wedge [.:;! ?] \$$

As we can see, if this above shown strategy is incorporated at the tokenization level, the negation problem is relatively well managed during the tokenization process. The authors provide some examples [20], as follows:

Sentence: *I don’t think I will enjoy it: it might be too spicy.*

*Tokenization*:

i  
don’t  
think\_NEG  
i\_NEG  
will\_NEG  
enjoy\_NEG  
it\_NEG  
:  
it  
might  
be  
too  
spicy  
.

In this latter case, even long-distance effects have been managed. In this sub-section we have seen how negation can be managed, either accompany by the rules defined in the previous sub-section III-B or in isolation as part of a smart tokenization strategy. In our approach, we have chosen to apply the *smart tokenization* strategy.

### C. Component 3: Fuzzy Sets approach to the SA Problem

In this sub-section we address the rest of components necessary to be able classify sentences into Positive or Negative and in addition qualify the strength of the associated polarity. In order to do that we must address the following topics:

- Describe and construct the *fuzzy sets* that will be utilised in this effort
- Describe the linguistic quantifiers or variables that will be implemented
- Provide the logic necessary -in combination with the lexicon and the fuzzy sets already mentioned- to address the sentiment classification problem at hand. This would include all necessary rules that must be devised from the NLP standpoint

- Describe the mechanics behind the whole process as we incorporate the use of fuzzy sets

1) *Linguistic Variables*: When deciding which *linguistic variables* to use, we came to the realisation that in our case it was needed to be a combination of accuracy and flexibility, so that a fair comparison with other classification methods could be easily established, and a degree-of polarity of a given sentence could be established. A linguistic modifier has the ability to change the level of truth values or the strength of a quantifier, and consequently it allows for the creation of new quantifiers words that provide a greater granularity for classification than using the mentioned quantifier in isolation. In our research, the classification labels for the **intensity** of semantic orientation and positive/negative polarity of a given sentence are identified as:

- Poor
- Slightly
- Moderate
- Very
- Most

According to G.A. Miller [22], *7 plus or minus 2*, is the effective number of categories that a subject (individual or person) can handle. We have chosen a conservative approach and have devised 5 labels (7 minus 2), symmetrically distributed in the interval [0...1]; our choice of trapezoidal function obeys to the fact that it generalises a triangular function and we have aimed for more than one value at the top of every category. We either classify a sentence as *Objective* or *Subjective*. Then, if it is considered to be *Subjective*, it could either be *Negative* or *Positive*, and then again either *Very Positive* or *Positive*, or *Very Negative* or *Negative*, etc. (see Fig. 1).

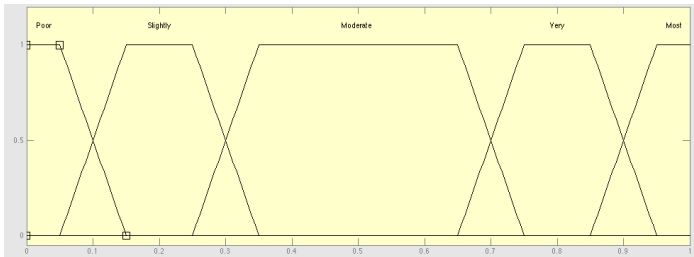


Fig. 1. Linguistic variables and membership functions (MF) using a trapezoidal function

A generic trapezoidal membership function would take the form:

$$\mu_{\bar{A}}(x) = \begin{cases} 0 & \text{if } x \leq a; \\ \frac{(x-a)}{(b-a)} & \text{if } a \leq x \leq b; \\ 1 & \text{if } b \leq x \leq c; \\ \frac{(d-x)}{(d-c)} & \text{if } c \leq x \leq d; \\ 0 & \text{if } d \leq x. \end{cases} \quad (1)$$

A trapezoidal membership function (MF) can be represented using the following 4-tuple  $(a, b, c, d)$ . These are

the specifics of the MFs functions devised for our hybrid model.

- MF (Poor): (0, 0, 0.050, 0.150)
- MF (Slightly): (0.050, 0.150, 0.250, 0.350)
- MF (Moderate): (0.250, 0.350, 0.650, 0.750)
- MF (Very): (0.650, 0.750, 0.850, 0.950)
- MF (Most): (0.850, 0.950, 1, 1)

2) *The mechanics of the fuzzy sets method applied to the SA problem*: In sub-section III-D we address the mechanics of how the opinion lexicon combined with the Semantic Rules and the fuzzy sets produce a fuzzy method to address the SA problem. First of all, let us present a diagram for the Hybrid Proposed model (Fig. 2). We do rely heavily on the off-line use of SentiWordNet [23, 15, 16, 24].

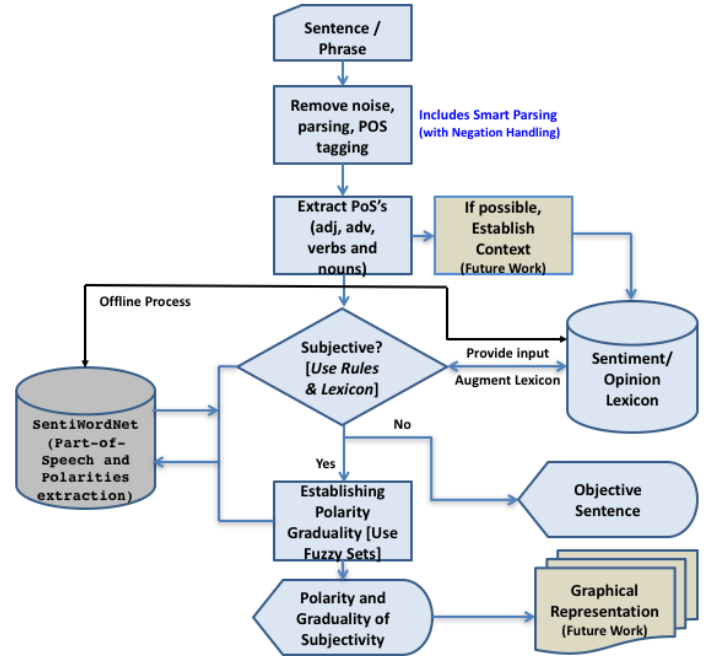


Fig. 2. View of our Hybrid proposed approach

#### D. Semantic Orientation (SO) Calculation and Granularity Improvements in Polarity Intensity

Our approach to determining the semantic orientation (SO) of the sentences belonging in our test dataset, takes a 2-step process, where Step 2 inherits the output of Step 1:

- 1) **Hybrid Standard Classification (HSC)**: this process is performed by using a number of techniques, as follows. Every intermediate step has as an outcome a list of features as specified:
  - a) Run tokenization & tagging process
  - b) Takes the output from the tokenization, misspelling errors cleansing, PoS tagging and smart parsing processes
  - c) Applying some of the Semantic Rules (presented in Table I), with

- emphasis on managing special particles  $\{unless, despite, but, while, however\}$  and the application of negation-handling through *smart parsing*. This step would imply changing the polarity of a given word-sentiment-carrying particle if such a particle is *negated*. It generates a list with the sentence re-structured as per the semantic rules applied
- d) Extract essential particles that convey sentiment/opinion (adjective, nouns, verbs and adverbs). It generates a list with the key particles and associated words
  - e) Per each sentence, it searches the Opinion Lexicon and substitute the words in the expression with the associated match in the lexicon
  - f) An exception List is generated for those key-words that are not in the Opinion Lexicon
  - g) The semantic orientation of each sentence is calculated by taking into consideration:
    - \* the word-label in the Lexicon (POS, NEG or OBJ)
    - \* the Positive & Negative Scores of the word in the opinion lexicon
 The calculations are performed and a classification label of *POS/NEG/OBJ/NOSOR* is assigned to each sentence. These features are returned in a list format.
  - h) The resulting list is re-scanned and for each sentence, those words labelled as OBJ are identified and a second-pass is performed to determine whether the words initially labelled as OBJ words carry a POS or a NEG sentiment (those words were initially marked as OBJ, because either no Positive/Negative score was found in the Lexicon or no POS/NEG label was found in the Lexicon). This situation is mostly related to relaying on a Sentiment Lexicon that *still* must/will grow and mature. We deal with this group of words by using the services of a dictionary previously generated for all sentences being processed. Then, the words are compared against the polarity of the results of the first pass of our *hybrid classifier*, and the *final polarity* in this specific context is determined by choosing the polarity that occurs more often for sentences where the considered particle/word appears. Additionally, the words labelled as NOSOR (no semantic orientation) are identified. This group will constitute an exception as the label is generated when none of the particles isolated in a given sentence is present in the Opinion Lexicon.
  - i) The new list with the classification POS/NEG of all sentences in the test dataset is generated substituting the particles initially marked as

OBJ, with their final label, either POS or NEG (exception to this is the case of the sentences labelled as NOSOR; for the latter, no classification can be offered at this point).

- 2) **Hybrid Advanced Classification (HAC):** this approach enhances the standard classification process by incorporating:

- Determination of the degree with which a given sentence lean towards being *positive* or *negative* (a fuzzy approach rather than a crisp method): for this, we will use the material presented in subsection III-C. Let us assume that *IP* corresponds to the Intensity of Polarity (Polarity Scores) of a given word,  $Word_i$

$$\min \left( IP(Word_1) \dots IP(Word_n) \right) = \delta, \quad (2)$$

where  $Word_i$  is the  $i^{esm}$  sentiment-carrying word in a given sentence, and  $i = 1 \dots numwords$ , where  $numwords$  corresponds to the total number of sentiment-conveying words,  $n$ , found in the Opinion Lexicon for a given sentence  $S_k$ ;  $\delta \in [0 \dots 1]$ .

$$\max_j \left( \mu_j(\delta) \right) = \beta, \quad (3)$$

where,  $\mu_j$ , with  $j$  taking all the possible 5 values of *poor, slightly, moderate, very, most*, corresponds to the evaluation of the membership function  $\mu_j$  for the sub-index  $j$ ;  $\beta \in [0 \dots 1]$ . In essence, we calculate the T-conorm represented by the value  $\beta$ .

- Diagnosing when a given sentence could be considered rather *objective/neutral* as opposed to either *positive* or *negative*

Let us calculate the polarity of a given sentence  $X$  in order to illustrate the mechanism used. In the example sentence below, only five words are opinion-conveying words and four of them are available in the sentiment lexicon. Hence, these four words will be the ones used to determine the polarity of the example-sentence  $X$ . Notice below the actual instances of the related words in the sentiment lexicon, showing the associated part-of-speech term and positive, negative and objective scores.

Sentence: "Many good actors. However, the film was simplistic, silly, unrealistic and tedious".

Sentiment-conveying words:

- Good: (#(good s 0.75 0.0 0.25 0 0) pos)
- Simplistic: (#(simplistic s 0.0 0.0 1.0 0 0) neg)
- Tedious: (#(tedious nopostag nonsc nopsc nocobj 0) neg)
- Silly: (#(silly n 0.125 0.0 0.875 0 0) neg)

In this particular example, our algorithm will look at the sentence and will notice the presence of the particle *however* and will follow the associated rule (Rule 15) in Table I. As such, the sub-sentence previous to the particle *however* will be discarded as contributing to the semantic

orientation of the sentence being analysed, and only the second sub-sentence, right after the particle *however* will be utilised (the words *simplistic*, *tedious*, *unrealistic* and *silly*).

For those cases in which our Sentiment Lexicon does not count with any of terms included in a specific sentence, our hybrid could not produce a semantic orientation output, as the *lexicon* is the key to our process. As an alternative, we have incorporated a process in which we build a dictionary based on all sentences in the dataset. A semantic orientation is then assigned based on a count of frequency or specific words that have appeared previously and have been labelled as Positive/Negative. The full description of this process is out of the scope of this article.

#### E. Experimental Results

Our Hybrid Method utilises a number of different components, coming from different disciplines, to achieve its results. This outcome, presented in Table III, are very encouraging. When we incorporate the fuzzy set approach (HAC), performance indicators are the same as in HSC, but we can add a much better granularity level in the classification process. See Table IV and Table V for details.

TABLE III  
HYBRID HSC/HAC CLASSIFIER - Key Performance Indexes

Indicator	Score
Accuracy	0.76
Precision	0.73
Recall	0.83
F1	0.77

TABLE IV  
HAC CLASSIFIER RESULTS FOR POS DATASET - Increased Granularity

False Negatives	929
True Positives	4,402
Poorly	577
Slightly	1,106
Moderate	1,041
Very	1,365
Most	313
Total Number of Sentences	5,331

#### IV. CONCLUSIONS

In general, our proposed hybrid system works very well with a high level of *accuracy* and *precision*. Indeed, the fact that our hybrid system improved the results obtained when we applied Naïve Bayes (NB) and Maximum Entropy (ME) to the same dataset **satisfies** one of our initial hypothesis, that a hybrid method using natural language processing techniques, semantic rules and

TABLE V  
HAC CLASSIFIER RESULTS FOR NEG DATASET - Increased Granularity

False Positives	1,646
True Negatives	3,685
Poorly	770
Slightly	1,089
Moderate	789
Very	864
Most	173
Total Number of Sentences	5,331

fuzzy sets should be able to perform well. Another benefit of our proposed system is that we have managed to identify different strengths in the polarity degree of the input sentences with regard to a specific base-case (negative or positive). By the utilisation of fuzzy sets we can determine when a given sentence has a stronger/weaker intensity in terms of polarity than another one in the dataset. Our expectation is that the quality of the content of SentiWordNet should continue improving with time. Those enhancements will contribute to the betterment of our proposed hybrid solution as it will reflect positively in the quality of our Opinion Lexicon. In theory, as time passes, both, SentiWordNet and our proposed Opinion Lexicon should become better and more complete. In essence, hybrid techniques can play an important role in the advancement of the Sentiment/Opinion Analysis discipline by combining together the elements we described in our original research hypothesis.

#### REFERENCES

- [1] M. D. Cock, U. Bodenhofer, and E. E. Kerre, "Modelling linguistic expressions using fuzzy relations," in *Proceedings of the 6th International Conference on Soft Computing*, October 2000, pp. 353–360.
- [2] F. Dzogang, M.-J. Lesot, M. Rifqi, and B. Bouchon-Meunier, "Expressions of Graduality for Sentiments Analysis – A Survey," in *Fuzzy Systems (FUZZ)*, 2010 IEEE International Conference on, 2010, pp. 1–7.
- [3] B. Liu, *Sentiment Analysis and Opinion Mining*, 1st ed. Morgan and Claypool Publishers: Synthesis Lectures on Human Language Technologies, 2012.
- [4] O. Appel, F. Chiclana, and J. Carter, "Main Concepts, State of the Art and Future research Questions in Sentiment Analysis," *Acta Polytechnica Hungarica - Journal of Applied Sciences*, vol. 12, no. 3, pp. 87–108, May/June 2015. [Online]. Available: <http://dx.doi.org/10.12700/APH.12.3.2015.3.6>
- [5] B. Pang and L. Lee, *Opinion mining and sentiment analysis*. NOW: the essence of knowledge, Foundations and Trends in Information Retrieval, Vol. 2, Nos. 1–2, pp. 1–135, 2008.

- [6] —, “Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales,” in Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL ’05), ser. ACL’05, 2005, pp. 115–124.
- [7] —, “A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts,” Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, (ACL ’04), pp. 271–278, 2004.
- [8] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques,” Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing (EMNLP), vol. 10, pp. 79–86, 2002.
- [9] M. Sadegh and R. I. Z. A. Othman, “Combining lexicon-based and learning-based methods for twitter sentiment analysis,” International Journal of Computers & Technology, vol. 2, no. 3, pp. 171–178, June 2012.
- [10] M. Hu and B. Liu, “Mining and summarizing customer reviews,” Proceedings – ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004 full paper), Seattle, Washington, USA, Aug. 22–25, 2004.
- [11] V. Hatzivassiloglou and K. McKeown, “Predicting the semantic orientation of adjectives,” in Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL. New Brunswick, NJ, USA: ACL, 1997, pp. 174–181.
- [12] J. KAMPS, M. Marx, R. J. Mokken, and M. de Rijke, “Using WordNet to measure semantic orientations of adjectives,” in Proceedings of LREC-04, 4th International Conference on Language Resources and Evaluation, ser. LREC ’04, vol. IV, 2004, pp. 1115–1118.
- [13] J. Wiebe, “Learning subjective adjectives from corpora,” in Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence. AAAI Press, 2000, pp. 735–740. [Online]. Available: <http://dl.acm.org/citation.cfm?id=647288.721121>
- [14] V. Hatzivassiloglou and K. McKeown, “Towards the automatic identification of adjectival scales: Clustering adjectives according to meaning,” in ACL: 31st Annual Meeting of the Association for Computational Linguistics, 22-26 June 1993, Ohio State University, Columbus, Ohio, USA, Proceedings, L. K. Schubert, Ed. ACL, 1993, pp. 172–182.
- [15] A. Esuli and F. Sebastiani, “SentiWordNet – A Publicly Available Lexical Resource for Opinion Mining,” In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC06), pp. 417–422, 2006.
- [16] —, “Determining the semantic orientation of terms through gloss classification,” in Proceedings of the 14th ACM International Conference on Information and Knowledge Management, ser. CIKM ’05. New York, NY, USA: ACM, 2005, pp. 617–624. [Online]. Available: <http://doi.acm.org/10.1145/1099554.1099713>
- [17] S. Nadali, M. Murad, and R. Kadir, “Sentiment classification of customer reviews based on fuzzy logic,” In Information Technology (ITSim), 2010 International Symposium in (Volume:2), Kuala Lumpur, Malaysia, vol. 2, pp. 1037–1040, June 2010.
- [18] P. Subasic and A. Huettner, “Affect Analysis of Text Using Fuzzy Semantic Typing,” IEEE Transactions on Fuzzy Systems, vol. 9, no. 4, pp. 483–496, August 2001.
- [19] Y. Xie, Z. Chen, K. Zhang, Y. Cheng, D. K. Honbo, A. Agrawal, and A. N. Choudhary, “MuSES: a multilingual sentiment elicitation system for Social Media Data,” Intelligent Systems, IEEE, vol. 29, no. 4, pp. 34–42, July 2014.
- [20] C. Potts, “Sentiment Symposium Tutorial: Linguistic structure (part of the Sentiment Analysis Symposium held at San Francisco, november 8-9, 2011). Stanford Department of Linguistics, Stanford University. Accessed date: December 2014,” November 2011. [Online]. Available: <http://sentiment.christopherpotts.net/lingstruc.html>
- [21] S. R. Das, M. Y. Chen, T. V. Agarwal, C. Brooks, Y. shee Chan, D. Gibson, D. Leinweber, A. Martinez-Jerez, P. Raghuram, S. Rajagopalan, A. Ranade, M. Rubinstein, and P. Tufano, “Yahoo! for amazon: Sentiment extraction from small talk on the web,” in 8th Asia Pacific Finance Association Annual Conference, 2001.
- [22] G. Miller, “The magical number seven, plus or minus two: Some limits on our capacity for processing information,” Psychological Rev., vol. 63, pp. 81–97, 1956.
- [23] A. Esuli and F. Sebastiani, “SentiWordNet: a high-coverage lexical resource for opinion mining,” Institute of Information Science and Technologies (ISTI) of the Italian National Research Council (CNR), Tech. Rep. ISTI-PP-002/2007, Oct 2006. [Online]. Available: <http://tcc.itc.it/projects/ontotext/Publications/sentiWN-TR.pdf>
- [24] V. Singh, R. Piryani, A. Uddin, P. Waila, and Marisha, “Sentiment analysis of textual reviews; evaluating machine learning, unsupervised and sentiwordnet approaches,” in Knowledge and Smart Technology (KST), 2013 5th International Conference on, Jan 2013, pp. 122–127.