

A consensus approach to the sentiment analysis problem driven by support-based IOWA majority

Orestes Appel

Bissett School of Business, Mount Royal University, Calgary, Canada

Centre for Computational Intelligence, De Montfort University, Leicester, UK

Email: oappel@mtroyal.ca; orestes.appel@my365.dmu.ac.uk

Francisco Chiclana*, Jenny Carter

Centre for Computational Intelligence, De Montfort University, Leicester, UK

Email: chiclana@dmu.ac.uk; jennyc@dmu.ac.uk

Hamido Fujita

Iwate Prefectural University, Takizawa, Iwate, Japan

Email: issam@iwate-pu.ac.jp

January 5, 2017

Abstract

In group decision-making there are many situations where the opinion of the majority of participants is critical. The scenarios could be multiple, like a number of doctors finding commonality on the diagnose of an illness or parliament members looking for consensus on an specific law being passed. In this article we present a method that utilises Induced Ordered Weighted Averaging (IOWA) operators to aggregate a majority opinion from a number of Sentiment Analysis (SA) classification systems, where the latter occupy the role usually taken by human decision-makers as typically seen in group decision situations. In this case, the numerical outputs of different SA classification methods are used as input to a specific IOWA operator that is semantically close to the fuzzy linguistic quantifier ‘most of’. The object of the aggregation will be the intensity of the previously determined sentence polarity in such a way that the results represents what the majority think. During the experimental phase, the use of the IOWA operator

*Corresponding author

coupled with the linguistic quantifier ‘most’ ($IOWA_{most}$) proved to yield superior results compared to those achieved when utilising other techniques commonly applied when some sort of averaging is needed, such as arithmetic mean or median techniques.

Keywords: Sentiment Analysis; Hybrid Sentiment Analysis Method; Naïve Bayes; Maximum Entropy; Consensus; Majority Support; Sentiment Aggregation; IOWA.

1 Introduction

Group decision making is a task where a number of agents get involved in a decision process to generate a value that represents their individual decisions in the group process.¹⁴ In the case of the Sentiment Analysis (SA) problem research effort presented in this article, the agents would be any number n of SA classification methods, where $n \geq 2$. Experiments have been conducted using three methods: (a) Naïve Bayes,²³ (b) Maximum Entropy,³ and (c) a Hybrid Approach to the SA problem devised by the authors.²

In the current article, additional insights are provided with the aim of continuing to improve the results obtained in our previously published research addressing the Sentiment Analysis problem.² As discussed in the previous paragraph, the central idea in this article is that several classification methods could be utilised in such a way that each of them performs the classification task following their intrinsic characteristics and design principles. Then, an appropriate model should be put in place to account in a sensible manner for the opinions of all of them. We would like to obtain a classification value that articulates all of them, that summarises the collective opinion of these methods, with the caveat that we would like the final classification value *to reflect the opinion of the majority*. In particular, we would like to compute what the opinion of the majority is with respect to the intensity of the polarity of a given sentence.

The concept of aggregating diverse methods recommendations is not technique dependent. Nevertheless, a proper aggregation method must be chosen. *Arithmetic mean* and *median* are central tendency values/techniques that have been used in the past.²⁹ However, we are in search of an aggregation mechanism that gives more importance to the classification output of some methods depending on the characteristics of the values they produce. This can be achieved using aggregation operators based on Yager’s Ordered Weighted Averaging (OWA) operator.⁴² The OWA operator will be discussed further to present our rationale for having selected an Induced Ordered Weighted Averaging (IOWA) operator^{10,44} as the aggregation mechanism of the outcome of several opinion classification methods using an induced guiding principle. In order to test our ideas, we will utilise the classification outcomes of Naïve Bayes, Maximum Entropy and the Hybrid Advanced Classification method presented in² as the individual polarity classification values to derive a consensus IOWA majority based polarity classification for the SA problem.

For a summarised survey on SA, please refer to the article by Appel et al.¹ For a complete review of the evolution of the SA field, please consult the very thorough work of Ravi and Ravi.³¹ For a focused account on recent advances in SA techniques, please access the works by Cambria et al.,⁷ while for *semantic analysis* the reader is referred to the work of Cambria et al.⁸ For a complete review of the IOWA operator topic, please refer to the work of Yager and Filev.⁴⁴ For a full review of the utilisation of OWA operators in aggregation processes in multi-criteria decision-making, access.⁴² Peláez et al.²⁵ provide an analysis of OWA operators in decision-making when the objective is to model the majority concept.

The rest of this paper is organised as follows: Section 2 presents related work done using OWA based operators, whilst Section 3 covers the basic concepts and properties of OWA and IOWA operators. Section 4 addresses the role of IOWA operators and fuzzy majority in collective decision-making; in order to provide context, Section 5 covers the hybrid method introduced by the authors in,² as the approach presented in this article represents an enhancement to this method² in terms of obtaining a majority sentiment classification opinion. Section 6 summarises the proposed IOWA approach to Sentiment Analysis in order to model consensus. Section 7 covers the experimental results obtained when applying the new majority based proposed methodology introduced in section 6. Section 8 closes the paper with some conclusions and the mention of further work to be conducted.

2 Related work

A number of authors have explored the utilisation of members of the family of OWA operators in different situations and domains, with Pasi and Yager providing comprehensive information about OWA operators.²⁴ Boroushaki and Malczewski⁵ present a very interesting example of applying a fuzzy majority approach for Geographical Information Systems based on multi-criteria group decision-making. Bordogna and Sterlacchini⁴ address a multi-criteria group decision making process based on the soft fusion of coherent evaluations of spatial alternatives (GIS-Spatial Analysis). In the work of Wei and Yuan³³ we can learn about the application to coal mine safety of linguistic aggregation operators in order to achieve effective decision-making. Mata et al.¹⁷ present the utilisation of a Type-1 OWA operator⁴⁷ as a vehicle to obtain aggregation in the presence of unbalanced fuzzy linguistic information in decision making problems, while Chiclana and Zhou¹¹ demonstrate that type-2 fuzzy sets can be effectively defuzzified using a Type-1 OWA alpha-level aggregation approach.⁴⁸

As we know, in multiple attribute decision-making situations, optimistic and pessimistic extremes are represented by maximum and minimum. Wei et al.³² propose a method based on IOWA operators in multiple attribute decision-making, in order to capture human attitudes that fall between the two extremes points

of optimism and pessimism. In,³⁰ Qian and Xu extend the properties of IOWA operators by incorporating linguistic preference information in applications in group decision making. Mata et al.¹⁸ propose a Type-1 OWA methodology devised to achieve consensus in multi-granular linguistic contexts. The work of Peláez et al.²⁵ on OWA operators in decision-making aimed to obtaining the opinion of the majority is very influential. More recently, Yager and Alajlan⁴⁰ addressed again the problem of obtaining a consensus subjective probability distribution from the individual opinions of a group of agents about the subjective probability distribution. In,⁴¹ Yager and Alajlan revise the parameterization aspects of OWA aggregation operators. The authors stress the fact that the aforementioned parameterization is achieved by the characterizing OWA weights. Yager and Alajlan expand on a number of different paths to provide these characterizing OWA weights. As typically the importance of the values being aggregated is application-dependent and the arguments have different importances, it becomes key “appropriately combining the individual argument weights with the characterizing weights of the operator to obtain operational weights to be used in the actual aggregation”.⁴¹ The authors present “the use of a vector containing the prescribed weights and the use of a function called the weight generating function from which the characterizing can be extracted”.⁴¹ Finally, it is worth mentioning Perkins’s work²⁹ on the use of median, voting and arithmetic mean when aggregating multiple classification results in SA, as it is relevant to the present article.

3 IOWA Operators

Usually, the first step of a group decision making resolution process is that of aggregating the information from which to derive a group solution to the problem. Yager’s OWA operator⁴² has been proved to be extremely useful in these decision making problems because it allows to implement the concept of *fuzzy majority*.³⁸

Definition 1 (OWA Operator). *An OWA operator of dimension n is a function $F: \mathbb{R}^n \rightarrow \mathbb{R}$, that has associated a set of weights or weighting vector $W = (w_1, \dots, w_n)$ to it, so that $w_i \in [0, 1]$ and $\sum_{i=1}^n w_i = 1$, and is defined to aggregate a list of real values $\{a_1, \dots, a_n\}$ according to the following expression:*

$$F(a_1, \dots, a_n) = \sum_{i=1}^n w_i \cdot a_{\sigma(i)},$$

being $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ a permutation such that $a_{\sigma(i)} \geq a_{\sigma(i+1)}$, $\forall i = 1, \dots, n-1$, i.e., $a_{\sigma(i)}$ is the i -th highest value in the set $\{a_1, \dots, a_n\}$.

If B is the vector whose components are the ordered arguments values, $b_i = a_{\sigma(i)}$, then:

$$F(a_1, a_2, \dots, a_n) = W^T B \quad (1)$$

OWA operators are one of the most commonly used operators in multi-criteria decision making and aggregation in situations where only some portion of the criteria must be satisfied³⁸. However, as we will present in Section 4, the IOWA operator, a general type of OWA operator with a specific semantic in the aggregation process, provides a good representation of the *majority concept* in multi-criteria decision making processes.

Mitchell and Estrakh in¹⁹ described a modified OWA operator in which the input arguments are not re-arranged according to their values but rather using a function of the arguments. Inspired by this work, Yager and Filev introduced in⁴⁴ a more general type of OWA operator, which they named the Induced OWA (IOWA) operator:

Definition 2 (IOWA Operator). *An IOWA operator of dimension n is a mapping $I - F: (\mathbb{R} \times \mathbb{R})^n \rightarrow \mathbb{R}$, which has an associated set of weights $W = (w_1, \dots, w_n)$ to it, so that $w_i \in [0, 1]$, $\sum_{i=1}^n w_i = 1$,*

$$I - F(\langle u_1, a_1 \rangle, \dots, \langle u_n, a_n \rangle) = \sum_{i=1}^n w_i \cdot a_{\sigma(i)},$$

and $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ is a permutation function such that $u_{\sigma(i)} \geq u_{\sigma(i+1)}$, $\forall i = 1, \dots, n - 1$.

In the above definition the reordering of the set of values to aggregate, $\{a_1, \dots, a_n\}$, is induced by the reordering of the set of values $\{u_1, \dots, u_n\}$ associated to them, which is based upon their magnitude. Yager and Filev called the vector of values (u_1, \dots, u_n) the order inducing vector and $\{a_1, \dots, a_n\}$ the values of the argument variable.^{39, 44} Thus, the main difference between the OWA operator and the IOWA operator is the reordering step of the argument variable. In the case of OWA operator the reordering is based upon the magnitude of the values to be aggregated, while in the case of IOWA operator an order inducing vector is used as the criterion to induce that reordering of the values to aggregate. Obviously, an immediate consequence of definition 2 is that if the order inducing vector components coincide with the argument values then the IOWA operator reduces to the OWA operator. In fact, the OWA operator as well as the weighted average (WA) operator are included in the more general class of IOWA operators, which means that the IOWA operators allow to take control of the aggregation stage of any multi-criteria decision making problem in the sense that importance can be given to the magnitude of the values to be aggregated as the OWA operators do or to the information sources as the WA operators do. In fact, the IOWA operator, in our opinion, can play a

significant role in the proposed Hybrid Solution to the SA problem as elaborated in Section 6.

An issue in the definition of the OWA and IOWA operator is how to obtain the associated weighting vector. In,⁴² Yager proposed two ways to obtain W . The first approach is to use some kind of learning mechanism using some sample data; and the second approach is to try to give some semantics or meaning to the weights. The latter allowed applications in the area of quantifier guided aggregations³⁸ because weights are derived “from a functional form of the linguistic quantifier”.²⁴ This approach would be favourable for the problem we are considering.

According to Pasi and Yager, let $Q : [0, 1] \rightarrow [0, 1]$ be a function such that $Q(0) = 0$, $Q(1) = 1$, and $Q(x) \geq Q(y)$ for $x > y$ corresponding to a fuzzy set representation of a proportional monotone quantifier. Then, for a given value $x \in [0, 1]$, $Q(x)$ is the *degree* to which x *satisfies* the fuzzy concept being represented by the *quantifier*.²⁴ Based on function Q , the elements of the OWA weighting vector are determined in the following way:²⁴

$$w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \quad (2)$$

Hence, w_i represents the increase of satisfaction in getting i with respect to $(i - 1)$ criteria satisfied.

Some examples of linguistic quantifiers, depicted in Fig. 1, are “*at least half*”, “*most of*” and “*as many as possible*”, which can be represented by the following function

$$Q(r) = \begin{cases} 0 & \text{if } 0 \leq r < a \\ \frac{r-a}{b-a} & \text{if } a \leq r \leq b \\ 1 & \text{if } b < r \leq 1 \end{cases} \quad (3)$$

using the values $(0, 0.5)$, $(0.3, 0.8)$ and $(0.5, 1)$ for (a, b) , respectively.¹⁵

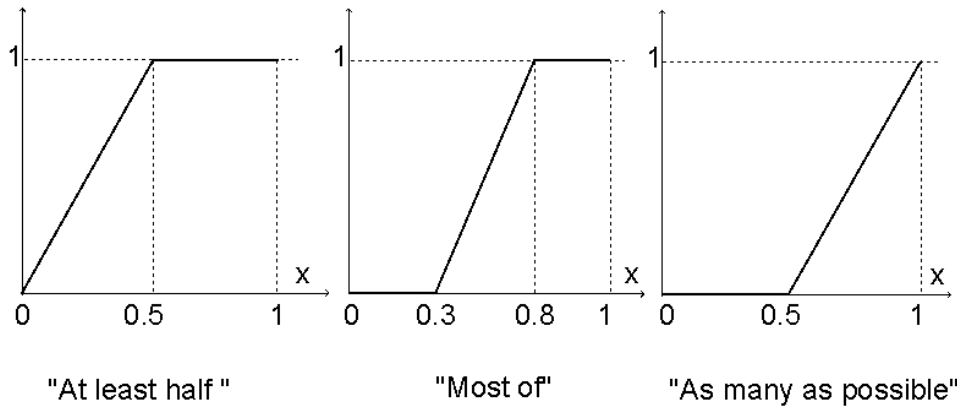


Figure 1: Linguistic quantifiers “*at least half*”, “*most of*” and “*as many as possible*”

Alternative representations for the concept of fuzzy majority can be found in the literature. For example,

Yager in³⁸ considered the parameterized family quantifiers $Q(r) = r^a$ ($a \geq 0$) for such representation. This family of functions guarantees that:¹⁰ (i) all the experts contribute to the final aggregated value (strict monotonicity property), and (ii) associates, when $a \in [0, 1]$, higher weight values to the aggregated values with associated higher importance values (concavity property).

4 Fuzzy Majority in Collective Decision Making modelled with an IOWA Operator

It has been already established by Yager^{37,44} that the OWA operators provide a *parameterized* family of mean type aggregation operators. The parameterized aspect is directly associated to the weighting vector. In this section we will take a closer look at OWA operators, fuzzy majority and other related decision making aspects.

4.1 The Linguistic Quantifier in Fuzzy Logic

In the same way as other fuzzy logic concepts relate to classical logic, the linguistic quantifier generalises the idea of quantification of classical logic. If we recall some articles from Zadeh, e.g.,^{45,46} in classical logic we deal with two types of quantifiers that can be used in propositions: the universal quantifier (*for all*) and the existential quantifier (*there exists*). According to Pasi and Yager,²⁴ by using linguistic quantifiers we are capable of referencing a variable number of elements of the domain of discourse. This referencing can be done in a *crisp* way or in a *vague* (fuzzy) manner as Table 1 shows. Pasi and Yager²⁴ also differentiate between two types of fuzzy quantified propositions as presented in Table 2.

Referencing type	Examples
Crisp	<i>at least k</i> of the elements, <i>half</i> of the elements, <i>all</i> of the elements
Vague (fuzzy)	<i>most</i> of the elements, <i>some</i> of the elements, <i>approximately k</i> of the elements

Table 1: Crisp and fuzzy referencing to elements of the domain of discourse

According to Zadeh,⁴⁵ in fuzzy logic the quantifiers have been defined as fuzzy subsets of two main types: absolute and proportional. As discussed in,²⁴ “absolute quantifiers, such as *about 7*, *almost 6*,” are modelled as fuzzy subsets of \mathbb{R} with membership function $Q(x)$ indicating the degree to which the amount x satisfies the concept Q . In addition, as per,²⁴ “proportional quantifiers like *most*, or *about 70%*,” as mentioned above are modelled as fuzzy subsets of the unit interval with $Q(x)$ indicating the degree to which the proportion x satisfies the concept Q .

Fuzzy quantified proposition	Components	Statement	Examples
Q X are Y	Q = Linguistic quantifier X = set of elements Y = a fuzzy predicate	Q elements of set X satisfy the fuzzy predicate Y	Most of the criteria are satisfied by alternative A_i Q = <i>most</i> X = set of <i>criteria</i> Y = <i>satisfies alternative A_i</i>
Q B X are Y	Q, X and Y as above B = fuzzy predicate	Q elements of set X which satisfy the fuzzy predicate B also satisfy the fuzzy predicate Y	Most of the important criteria are satisfied by alternative A_i Q, X and Y as above B = <i>important</i>

Table 2: Types of fuzzy quantified propositions

4.2 Linguistic quantifiers as soft specifications of majority-based aggregation

The focus of the discussion will be on monotonic non-decreasing linguistic quantifiers, like *most* and *at least*. We will concentrate of the quantifier “most” as we aim to model a *majority*, as per Pasi and Yager,²⁴ by using linguistic quantifiers “in guiding an aggregation process aimed at computing a value which synthesizes the majority of values to be aggregated”. Indeed, Pasi and Yager in²⁴ discuss extensively whether the result of aggregating a collection of values with a quantifier that corresponds to the concept of *majority* will be representative of the majority of values. The semantics behind the aggregation being performed is the key to reflecting the concept of a *majority*. As such, the two alternatives in terms of OWA semantics presented by Pasi and Yager²⁴ are: (a) OWA operators as an aggregation guided by ‘majority’ linguistic quantifiers, and (b) IOWA operators as drivers of a *majority opinion*.

- (a) The semantic of OWA operators is an aggregation guided by ‘majority’ linguistic quantifiers.²⁴ As it is well known, the weights of the OWA operator will determine the behaviour of the aggregation operator by “*emphasizing* or *demphasising* different components in the aggregation.”. As per,²⁴ one possible semantics of the OWA weights refers to the OWA operator as a generalisation of the idea of an averaging or summarising operator. (i.e. $w_i = 1/n$ for all i yields the simple average as all elements in the aggregation contribute equally to the final result), while another semantics to be considered refers to the OWA operator as a generalisation of the classical logic quantifiers *there exists* and *for all*. However, this second semantics is considered by Pasi and Yager as *not* really reflecting the concept of majority in group decision making applications. Hence, the authors suggest a different approach based on the IOWA operator.
- (b) Using IOWA operators to obtain a *majority opinion*. This corresponds to the concept of majority, as per the linguistic quantifier *most*, typically used in group decision making applications where more than one agent participates.²⁴ In this case, it is required an operator that calculates an average-like aggregation

of the “**majority** of values that are similar”. To achieve this, Pasi and Yager²⁴ propose an aggregation based on the utilisation of IOWA operator “with an inducing ordering variable which is based on a proximity metric over the elements to be aggregated,” which is based on the idea that “similar values must have close positions in the induced ordering in order to appropriately be aggregated”. This approach is elaborated below.

The final output of an IOWA operator should reflect in a closer manner the opinion of the majority if similar values are closer to each other in the induced vector.²⁴ Thus, what is needed is the ability to calculate the similarities between the values to be aggregated to compute “the values of the inducing variable of the IOWA operator”.²⁴ Such a function is defined using a binary support function, Sup , introduced by Yager in⁴³ where $Sup_\alpha(a, b)$ expresses the support from b for a at an α level of desired tolerance based on the premise of “the more similar two values are, the more they support each other”:

$$Sup_\alpha(a_i, a_j) = \begin{cases} 1 & \text{if } |a_i - a_j| < \alpha \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The higher the tolerance is, the less it is imposed that the two values have to be closer to each other to support each other. If we were to aggregate a set of values and we wanted to order them in increasing order of support, then for each value it is computed the sum of its support values with respect to the rest of values to be aggregated.²⁴ These overall supports are utilised as the values of the order inducing variable. In the event of having a tie in support between two values, a stricter tolerance level (decrease value of α) could be set to brake it. Thus, the use of an adequate support function will enable to induce an ordering based on proximity, which is key in understanding IOWA operators to generate a *majority-based aggregation* of the values to be aggregated via the linguistic quantifier *most* presented in Eq. (3) with values $(a, b) = (0.3, 0.8)$. Also, Pasi and Yager’s strategy implies that the construction of the weighting vector that appropriately implement the more *influence* in the aggregation result from the most supported individual values. Consequently the following process for the construction of the weighting vector from the induce support values is proposed:

1. Include in the definition of the overall support for a_i the similarity of the value a_i with itself:

$$t_i = s_i + 1.$$

2. On the basis of the t_i values, the weights of the weighting vector are computed as follow:

$$w_i = \frac{Q\left(\frac{t_i}{n}\right)}{\sum_{j=1}^n Q\left(\frac{t_j}{n}\right)} \quad (5)$$

“The value $Q(t_i/n)$ denotes the degree to which a given member of the considered set of values represents the *majority*” as per the linguistic quantifier Q .

As such, Eq. (5) is the weights semantic to apply in the proposed SA aggregation problem.

5 A Hybrid Approach to the SA Problem at the Sentence Level aimed to opinion consensus

In,² the authors describe a hybrid model for the SA (HSC/HAC) problem at the sentence level that is based on semantic rules, smart NLP techniques and fuzzy sets. The IOWA approach for aggregation that we are presenting in this article will be used to complement the aforementioned HSC/HAC model with the aim to arrive at a consensus sentiment classification opinion in SA.⁶ Before, the key concepts of the HSC/HAC approach to SA will be summarised, which will be followed by the methodology to using the IOWA operator described in this article ($IOWA_{most}^\alpha$) as an aid to augment the scope of its possible applications.

The HSC/HAC approach to SA works by:

1. Creating and utilising a sentiment lexicon with rich semantic attributes extracted from SentiWord-Net;^{12,13}
2. Using semantic rules paired with smart tokenisation and parsing to handle negation and other NLP aspects;
3. Implementing a fuzzy set methodology to model the intensity of the polarity of a given sentiment/opinion in a linguistic way.

In brief, the polarity of sentences is calculated and then the intensity of those polarities is computed. For completion, Fig. 2 presents the high-level diagram of the HSC/HAC model as described in.² Enhancements to the HSC/HAC approach to SA could come in the form of improvements in its accuracy, precision and recall. However, other enhancements might come in its specific applications to SA either in isolation or in combinations with other techniques. In this article, we address one of those specific applications, which is the scenario in which one might be interested in finding an aggregated sentiment polarity value representing the

opinion of the majority of approaches available to address the SA problem. In such a situation, as discussed before the IOWA aggregation mechanism is deemed as appropriate to integrate the outcomes of several SA approaches in deriving a consensus of opinion problem in SA, centred around obtaining the intensity of the sentence's polarity, just as it is carried out in decision-making research contexts. Fig. 3 depicts graphically the aggregation approach put forward in this article to achieve such consensus polarity, and the implementation of the proposed IOWA aggregation operator is elaborated in the next section.

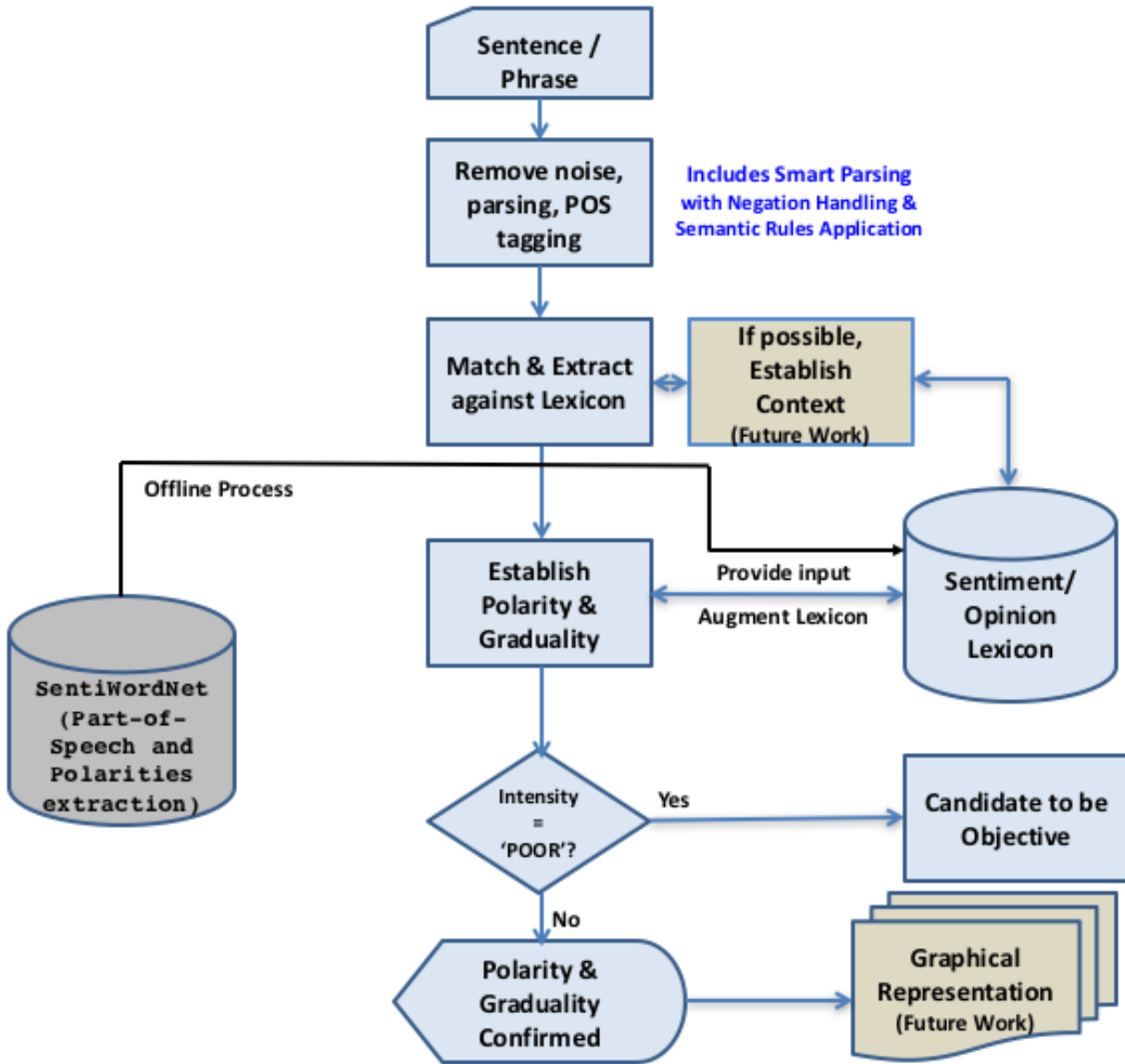


Figure 2: View of Hybrid Approach HSC/HAC

6 The Proposed IOWA Approach to Sentiment Analysis

Constructing a *majority opinion* could be explained as “...the collective evaluation of a majority of the agents involved in the decision problem”.²⁴ The following authors provide ample and detailed information about the OWA operators and its applications: León et al.¹⁶ and Pérez et al.;²⁶ Wu & Chiclana,³⁵ Wu et al.,³⁶ and Yager;⁴² Chiclana et al.⁹ and Pasi & Yager.²⁴ Of particular interest are the three latter ones. In addition, Bordogna & Sterlacchini⁴ and Boroushaki & Malczewski⁵ provide very good examples of real applications of OWA operators.

Let us for a moment consider the problem of determining subjectivity polarity for a given sentence S_k using the recommendations of several classification systems. In a way, each method to be used and applied to the aforementioned sentence S_k , can be seen as an ‘agent/person’ giving her opinion on whether the sentence S_k is positive or negative. In the end, we would like to collect all provided answers and summarise them with a sort of central tendency measure derived from the inputs received. Basically, in our context, and as discussed previously, we would like to *aggregate the polarity intensity* value of sentence S_k measured by using different classification methods. Hence, the final polarity value will be the ‘induced aggregation of the majority’ of the subjectivity intensity polarity of sentence S_k when one takes into consideration all the different contributions of all the participating classification methods.

The different applied classification methods will issue their individual *opinions* on whether a sentence is positive or negative, just as individual agents use their own judgement. We will denote these classification methods $\{M_1, M_2 \dots M_n\}$. These techniques will arrive to their respective conclusions using their own appraisal strategies. Each method will have their own peculiarities. For instance, the Naïve Bayes method will classify a bag-of-words features, the fuzzy method will look at the level of belonging of a given particle to a given category (fuzzy set), while the Maximum Entropy technique will apply its own methodology to verify the existence of one or a number of ‘opinion-carrying particles’ belonging in the predetermined part-of-speech. Fig. 3 shows a graphical representation of the way the IOWA operator with the semantic ‘most’ is implemented in order to achieve our pre-established objective.

In fact, we would like to think that in addition to obtaining the aggregation already targeted, we could in the future incorporate the *level of trust* that we have on each method, ensuring that well-established proven or accurate methods carry more weight than the rest as we would do when considering the opinions of a number of people, depending on how much we *trust* each of them. The hybrid method described in² could therefore be improved with the incorporation of an IOWA operator with the semantic of the linguistic quantifier *most*, namely, $IOWA_{most}$, to handle the numerical output of three classification methods: Naïve Bayes, Maximum Entropy and the HSC/HAC hybrid technique introduced in.²

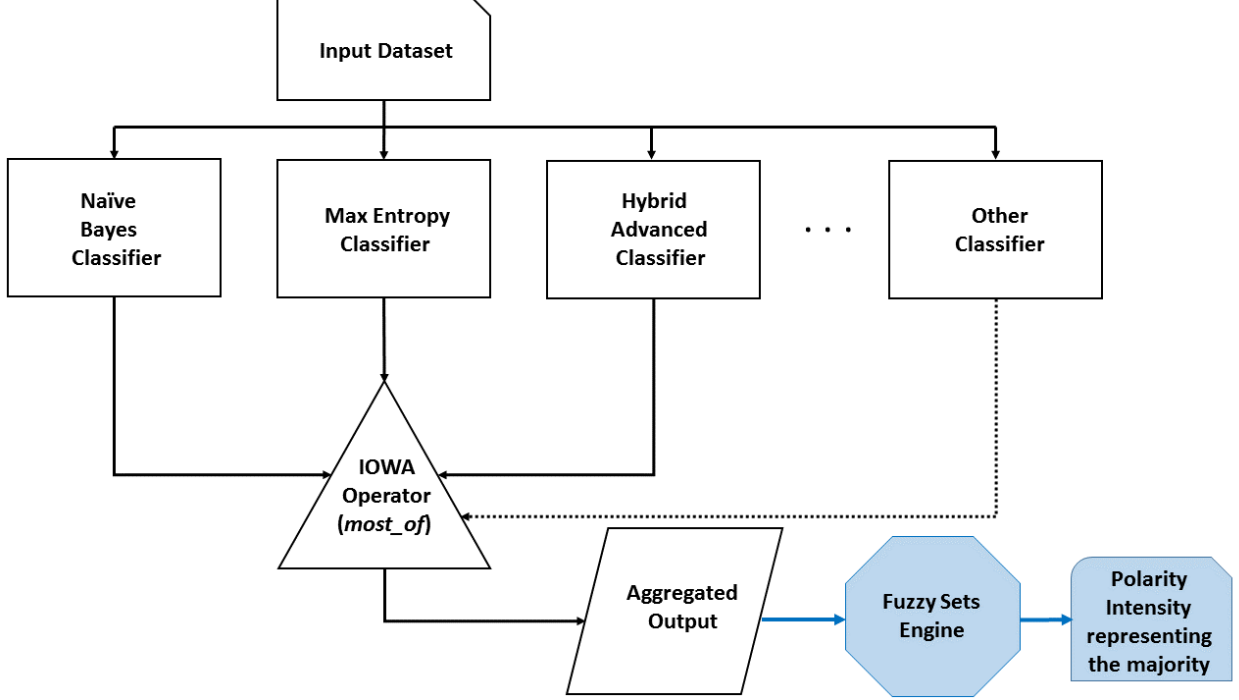


Figure 3: $IOWA_{most}$ operator aggregating classifier methods outputs

7 Experimental results obtained applying $IOWA_{most}^{\alpha}$ aggregation

In order to do a proper comparison, we will evaluate how the $IOWA_{most}^{\alpha}$ operator performs when compared to both *Arithmetic mean*²⁹ and *Median*.²⁹ However, we will firstly describe briefly the datasets utilised during the experimental phase.

7.1 Datasets used

As mentioned in Appel et al.,² Pang and Lee²² have published datasets that were utilised in SA experiments reported in.^{20,21,23} As such, it seems adequate to use the Movie Review Dataset provided by Pang and Lee (available at: <http://www.cs.cornell.edu/people/pabo/movie-review-data/>). In order to use the output of all classifiers as an input to the $IOWA_{most}^{\alpha}$ process all participating scores have been converted to the interval $[0, 1] \in \mathbb{R}$, where S_k corresponds to any sentence in the test dataset and $m_i = \{m_1, m_2, \dots, m_n\}$ represents the different classification methods i being aggregated ($n \geq 2$), then:

$$IOWA_{most}^{\alpha}[S_k](m_1, m_2, \dots, m_n) = \Theta^{S_k} \quad (6)$$

Once the aggregation with the semantic representing the *opinion of the majority* has been computed, the

intensity level to which the value Θ belongs is calculated. For that, we use the classification method presented in Appel et al.² that uses the following granules on the perception of the *intensity of the polarity*, either *positivity* or *negativity*, of a given sentence S : $G = \{Poor; Slight; Moderate; Very; Most\}$, represented using the following 4-tuple trapezoidal membership functions (MF),²⁸ as generally shown in Fig. 4:

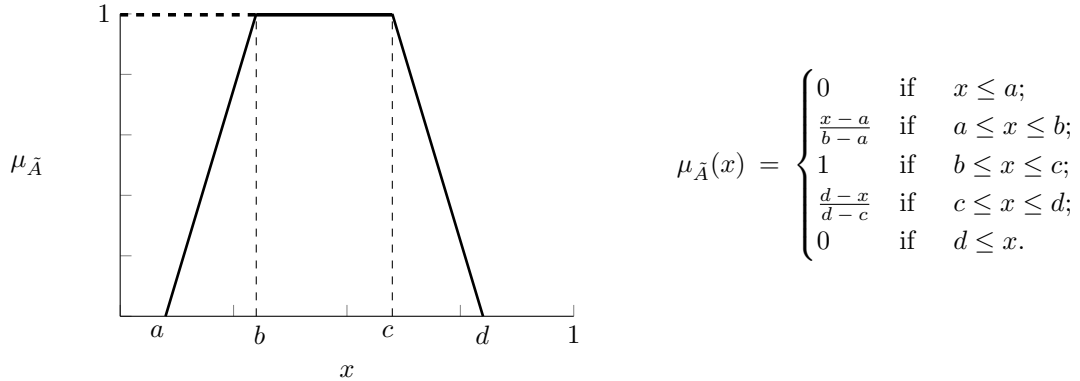


Figure 4: Trapezoidal membership function

- MF (Poor): (0, 0, 0.050, 0.150)
- MF (Slight): (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)

The aggregated value Θ^{S_k} previously computed in equation (6) will take on the value x in $\mu_{\bar{A}}(x)$ and in consequence a proper linguistic label belonging to G will be generated. This value represents the polarity intensity (how positive or how negative) of a given sentence S_k ($\mu_{\bar{A}}(\Theta^{S_k}) \in G$).

7.2 Comparison criteria

The comparison we are attempting to make might be challenging to achieve as the intensity of the polarity of a sentence would be vague in nature, as opposed to crisp. We are trying to see which method is semantically closer to the opinion of the majority among the participating methods. The classification of a sentence as belonging to one of the granules we defined in section 7.1, $G = \{Poor; Slight; Moderate; Very; Most\}$, is rather a subjective exercise. The datasets used count each (positive occurrences and negative occurrences) with 5,331 sentences. We have annotated 500 sentences, approximately 10%, assigning each of them a value

$v_k \in G$. These were estimated by looking at the classification outcomes of the three classifiers we are utilising as inputs and estimating a linguistic label in G that is representative of the opinion of the majority.

7.3 Non-OWA Aggregation - The outputs of the three classification methods combined using Median and Arithmetic Mean

Before we applied the $IOWA_{most}$ operator, we tested the idea of combining directly the results of the three chosen methods. The outcomes, which are summarised below, are not as good as those obtained by using the IOWA operator, as it will become evident later on in Section 7.4. This fact, basically, shows that the IOWA operator does a much better job at aggregating the individual outcomes of the three aforementioned techniques, by giving more weight to the leaning opinion of the majority. In essence, by properly weighting the advice of the three methods (NB, ME and the HSC/HAC approach) we do obtain a more realistic aggregation effect that represents the thinking of the majority.

The first aggregation method used for testing was based on the *Median*. Table 3 shows the associated Performance Indexes for *Median*.

Represents opinion of the majority	337
Does not represent opinion of the majority	163
% of success	67.40

Table 3: *Median* aggregation

The second aggregation method used for testing was based on the *Arithmetic mean*. Table 4 shows the associated Performance Indexes for *Arithmetic Mean*.

Represents opinion of the majority	388
Does not represent opinion of the majority	112
% of success	77.60

Table 4: *Arithmetic Mean* aggregation

7.4 IOWA Aggregation - The outputs of the three classification methods combined using the $IOWA_{most}^\alpha$ operator

The results of using $IOWA_{most}^\alpha$ with two different levels of tolerance ($\alpha = 0.3, 0.5$) are shown in Table 5 and Table 6.

The main difference between the results obtained when using different tolerance values (0.3 and 0.5) when $IOWA_{most}^\alpha$ is applied is not in whether the outcome will distance itself from representing the opinion of the majority, but rather to which *linguistic label* in G a specific sentence will be assigned. Depending on the

Represents opinion of the majority	500
Does not represent opinion of the majority	0
% of success	100.00

Table 5: $IOWA_{most}^{0.3}$ operator aggregation

Represents opinion of the majority	500
Does not represent opinion of the majority	0
% of success	100.00

Table 6: $IOWA_{most}^{0.5}$ operator aggregation

majority value calculated a sentence classified as ‘Moderate’ with a tolerance of 0.3 could now be labelled as ‘Very’ in terms of intensity, when the tolerance value changes to 0.5. In reality, the lower the *tolerance*, the more demanding the IOWA operator is on how closely the values in the aggregation support each other.

Table 7 presents a comparison of results with the outputs obtained using *Median* and *Arithmetic mean*. Notice that the aggregation results obtained using the IOWA operator are much more compelling than the rest. Especially, because $IOWA_{most}^{\alpha}$ *always* represents the targeted majority, and as a consequence it is the best option when compared to the other two methods tested. Therefore, these results simply re-enforce the applicability of carrying out the aggregation using the IOWA operator with a semantic associated to a specific linguistic quantifier (in this case, *most*).

Classification Method	Median	Arithmetic Mean	$IOWA_{most}^{tolerance=0.3}$	$IOWA_{most}^{tolerance=0.5}$
% of success	67.40%	77.60%	100.00%	100.00%

Table 7: The three aggregating methods - *Performance Indexes Compared*

7.5 Examples of applying the $IOWA_{most}$ operator to specific members of the dataset

This section presents examples of the application of the $IOWA_{most}^{\alpha}$ operator to several sentences. The samples include the output of three different classification methods, m_1, m_2, m_3 , where their outputs belong in the interval $[0, 1]$.

Example 1. $(m_1 \ m_2 \ m_3) = (0.959112 \ 0.500030 \ 1.00000)$.

Arithmetic mean = 0.819716

Median = 0.959112

IOWA (Tolerance = 0.50) = 0.819717

With a tolerance of 0.5, which does not enforce strict support among the values to be aggregated, all the elements contribute to the aggregation, generating a value that is extremely close to the arithmetic mean.

Example 2. $(m_1 m_2 m_3) = (0.564631 0.508914 1.000000)$.

Arithmetic mean = 0.691181

Median = 0.564631

IOWA (Tolerance = 0.50) = 0.536773

With a tolerance of 0.3, which does enforce a stricter support among the values to be aggregated (in this case the first two values), the value that represents the sentiment of the majority is closer to the elements with higher support: 0.564631 and 0.508914; as a consequence the generated value is not that close to the arithmetic mean, neither it is close to the median, although it is closer to the latter than the former.

Example 3. $(m_1 m_2 m_3) = (0.989550 0.682592 0.600000)$.

Arithmetic mean = 0.757380

Median = 0.682592

IOWA (Tolerance = 0.30) = 0.641296

With a tolerance of 0.3, again the IOWA aggregation with the semantic *most*, generates an aggregation between 0.682592 and 0.600000, which support each other, representing again the opinion of the majority.

7.6 The role of the tolerance parameter in the calculation of the support vector in IOWA

We have mentioned before that the IOWA operator used to generate the aggregation is formulated in such a way that a tolerance input is provided during the aggregation process. Let us look at the results obtained when we map the tolerance parameter against the polarity intensity classification of the test dataset.

The *tolerance* value (α) is used in the calculation of the support vector mentioned above in equation (4). The farther apart a_i and a_j , the higher the value of $|a_i - a_j|$. The higher the tolerance value α becomes, the more likely the support function $Sup_\alpha(a_i, a_j)$ will take a value of 1, which means that both values a_i and a_j will be part of the aggregation being performed. This translates into a situation in which: if a value a_i is too far apart from the rest of values a_j ($j \neq i$) so that its corresponding distances to the them exceed the defined tolerance for the support vector, $|a_i - a_j| > \alpha$, then that value contribution to the aggregation process will be very low as it will have assigned a value of $t_i = 1$. A tolerance value $\alpha = 1$ means that all values considered in the aggregation will contribute equally to the collective aggregated value regardless of how close they are to each other, and the $IOWA_{most}^1$ will coincide with the *arithmetic mean*. On the other hand, a tolerance value of, let us say, 0.1 will imply that the considered values for aggregation will have to be very close to each other to have a high support and higher contribution to the collective aggregated value. In our experiment, Fig. 5 shows that the highest the value of the *tolerance* parameter is, the more the intensity polarities are

distributed towards the linguistic labels ‘very’ and ‘most’, which are located towards the very right-end of the unit interval $[0, 1]$. Notice as well that the tipping point in the x-axis is the value 0.5 (in Fig. 5, *HAC* stands for Hybrid Advanced Classification, whilst *HACA* means Hybrid Advanced Classification with Aggregation, that are methods we introduced in²).

8 Conclusions and further work

Induced Ordered Weighted Averaging (IOWA) operators can certainly play a significant role in aggregating the opinions of a number of sentiment classification systems. The aforementioned operator works by producing a value that get significantly closer to the collective opinion of the participants. The $IOWA_{most}^\alpha$ used in this article conveys the semantic of the opinion of the majority, which is represented by the linguistic quantifier *most*. Its performance in identifying the *intensity of the opinion of the majority*, according to our experiments, surpassed the one exhibited by *Arithmetic Mean* and *Median* techniques. The results obtained are sensible as the $IOWA_{most}^\alpha$ operator produces results that gravitate towards the opinion of most of the input values being processed. In essence, $IOWA_{most}^\alpha$ produces a larger pull towards the values that support each other, driving the results in the direction of what the majority reflects.

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

- Investigate other OWA operators that could potentially produce a better aggregation representing the semantic *majority opinion*.
- In addition to obtaining the aggregation already mentioned among the classification methods, the level of trust that one would have on each method $\{M_1, M_2, \dots, M_n\}$ could be incorporated as well. A first approach to this issue could be based on providing to those better established, respected and proven methods more weight in the aggregation process, as one would do when pondering the opinions of a number of experts, depending on how much one *trusts* each of them.^{27, 34, 36}
- Utilise the OWA measure *Dispersion*, which calculates the degree to which all aggregates are used in the resulting final aggregation.²⁵ The idea would be to gain a deeper understanding on how the support vector is configured to contribute to the semantic of a majority opinion, depending on the data values participating in the aggregation.

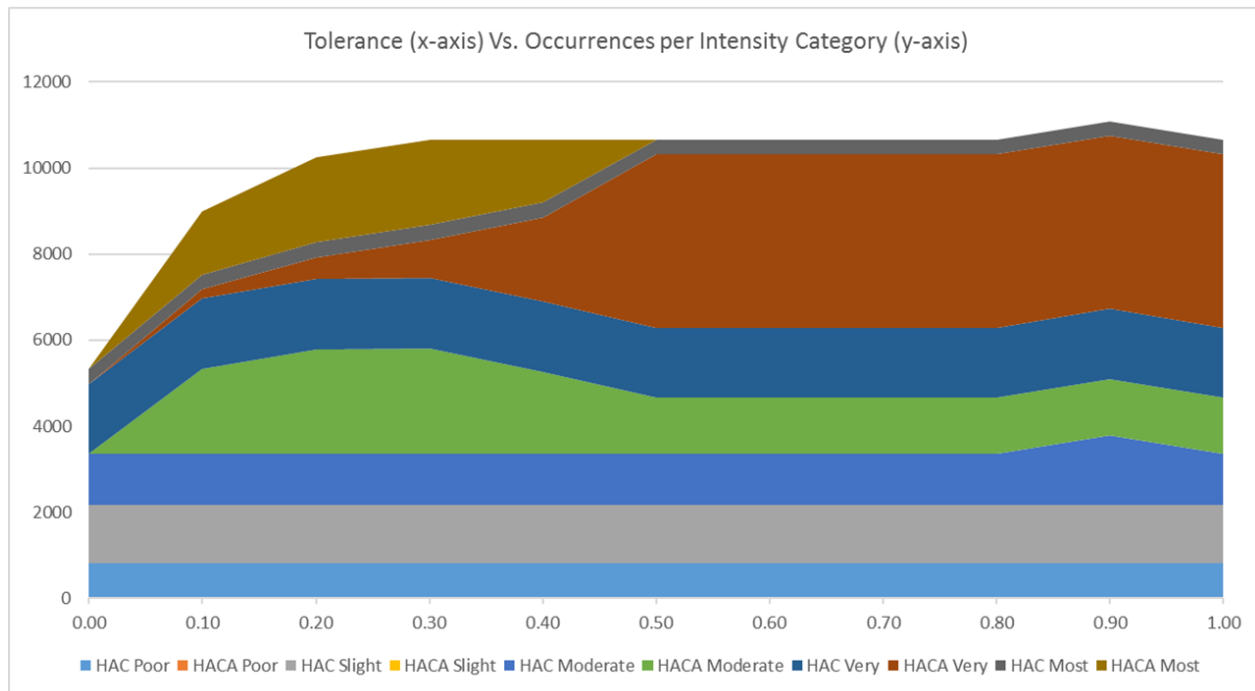


Figure 5: Tolerance Vs. Intensity of Polarity

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, May/June 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*, 2016, doi: 10.1016/j.knosys.2016.05.040.
- [3] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer Science + Business Media, LLC, 1st edition, 2006.
- [4] Gloria Bordogna and Simone Sterlacchini. A multi criteria group decision making process based on the soft fusion of coherent evaluations of spatial alternatives. In Lotfi A. Zadeh, Ali M. Abbasov, Ronald R. Yager, Shahnaz N. Shahbazova, and Marek Z. Reformat, editors, *Recent Developments and New Directions in Soft Computing*, volume 317 of *Studies in Fuzziness and Soft Computing*, pages 65–79. Springer International Publishing, 2014.

- [5] Soheil Boroushaki and Jacek Malczewski. Using the fuzzy majority approach for gis-based multicriteria group decision-making. *Comput. Geosci.*, 36(3):302–312, Mar 2010.
- [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, Bjorn Schuller, Yunqing Xia, and Catherine Havasi. New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, 28(2):15–21, March 2013.
- [8] Erik Cambria, Björn W. Schuller, Bing Liu, Haixun Wang, and Catherine Havasi. Knowledge-based approaches to concept-level sentiment analysis. *IEEE Intelligent Systems*, 28(2):12–14, 2013.
- [9] Francisco Chiclana, Francisco Herrera, and Enrique Herrera-Viedma. Integrating three representation models in fuzzy multipurpose decision making based on fuzzy preference relations. *Fuzzy Sets Syst.*, 97(1):33–48, Jul 1998.
- [10] Francisco Chiclana, Enrique Herrera-Viedma, Francisco Herrera, and Sergio Alonso. Some induced ordered weighted averaging operators and their use for solving group decision-making problems based on fuzzy preference relations. *European Journal of Operational Research*, 82(1):383–399, 2007.
- [11] Francisco Chiclana and Shang-Ming Zhou. Type-reduction of general type-2 fuzzy sets: the type-1 OWA approach. *International Journal of Intelligent Systems*, 28(5):505–522, 2013.
- [12] Andrea Esuli and Fabrizio Sebastiani. Determining the semantic orientation of terms through gloss classification. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM '05*, pages 617–624, New York, NY, USA, 2005. ACM.
- [13] Andrea Esuli and Fabrizio Sebastiani. SentiWordNet – A Publicly Available Lexical Resource for Opinion Mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC06)*, pages 417–422, 2006.
- [14] Francisco Herrera and Enrique Herrera-Viedma. Linguistic decision analysis: steps for solving decision problems under linguistic information. *Fuzzy Sets and Systems*, 115:67–82, 2000.
- [15] J. Kacprzyk. Group decision making with a fuzzy linguistic majority. *Fuzzy Sets and Systems*, 18:105–118, 1986.

- [16] Teresa León, Nuria Ramón, José L. Ruiz, and Inmaculada Sirvent. Using induced ordered weighted averaging (iowa) operators for aggregation in cross-efficiency evaluations. *International Journal of Intelligent Systems*, 29(12):1100–1116, 2014.
- [17] Francisco Mata, Luis G. Pérez, Francisco Chiclana, and Enrique Herrera-Viedma. Aggregation on Unbalanced Fuzzy Linguistic Information in Decision Problems based on Type-1 OWA Operator. *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE2015), Istanbul, August 2-5, 2015*, pages 1–6, August 2015.
- [18] Francisco Mata, Luis G. Pérez, Shang-Ming Zhou, and Francisco Chiclana. Type-1 OWA methodology to consensus reaching process in multi-granular linguistic contexts. *Knowledge-Based Systems*, 58:11–22, 2014.
- [19] H. B. Mitchell and D. D. Estrakh. A modified owa operator and its use in lossless dpcm image compression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5:429–436, 1997.
- [20] Bo Pang and Lillian Lee. 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)*, pages 271–278, 2004.
- [21] Bo Pang and Lillian Lee. 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)*, ACL'05, pages 115–124, 2005.
- [22] Bo Pang and Lillian 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.
- [23] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the Association for Computational Linguistics (ACL-02) Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 10:79–86, 2002.
- [24] Gabriella Pasi and Ronald R. Yager. Modeling the concept of majority opinion in group decision making. *Inf. Sci.*, 176(4):390–414, Feb 2006.
- [25] J.I. Peláez, J. M. Doña, and J.A. Gómez-Ruiz. Analysis of OWA operators in decision making for modelling the majority concept. *Applied Mathematics and Computation*, 186:1263–1275, 2007.

- [26] Luis G. Pérez, Francisco Mata, and Francisco Chiclana. Social Network Decision Making with Linguistic Trustworthiness-Based Induced OWA Operators. *International Journal of Intelligent Systems*, 29(12):1117–1137, 2014.
- [27] Luis G. Pérez, Francisco Mata, Francisco Chiclana, Gang Kou, and Enrique Herrera-Viedma. Modeling influence in group decision making. *Soft Computing*, 20(4):1653–1665, 2016.
- [28] Patrizia Perez-Asurmendi and Francisco Chiclana. Linguistic majorities with difference in support. *Applied Soft Computing*, 18:196–208, 2014.
- [29] Jacob Perkins. *Python Text Processing with NLTK 2.0 Cookbook*. Packt Publishing, 2010.
- [30] Gang Qian and Ze-Shui Xu. Extended IOWA operator and its application to group decision making with linguistic preference information. *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006*, pages 1662–1666, August 2006.
- [31] 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.
- [32] Chunfu Wei, Zheng Pei, and Bo Li. Multiple Attribute Decision Making Based on Induced OWA operator. *Proceedings of the 2009 IEEE International Conference on Fuzzy Systems (FUZZ- IEEE2009), Jeju Island, Korea, August 20-24, 2009*, pages 1763–1766, August 2009.
- [33] Chunfu Wei and Ruifu Yuan. Decision-making Method based on Linguistic Aggregation Operators for Coal Mine Safety Evaluation. In *2010 International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pages 17–20, Nov 2010.
- [34] Jian Wu and Francisco Chiclana. A social network analysis trust-consensus based approach to group decision-making problems with interval-valued fuzzy reciprocal preference relations. *Knowledge-Based Systems*, 59:97–107, 2014.
- [35] Jian Wu and Francisco Chiclana. Visual information feedback mechanism and attitudinal prioritisation method for group decision making with triangular fuzzy complementary preference relations. *Information Sciences*, 279:716–734, 2014.
- [36] 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.
- [37] Ronald R. Yager. Families of OWA Operators. *Fuzzy Sets Syst.*, 59(2):125–148, Oct 1993.

- [38] Ronald R. Yager. Quantifier guided aggregation using OWA operators. *International Journal of Intelligent Systems*, 11(1):49–73, 1996.
- [39] Ronald R. Yager. Induced aggregation operators. *Fuzzy Sets and Systems*, 137:59–69, 2003.
- [40] Ronald R. Yager and Naif Alajlan. An intelligent interactive approach to group aggregation of subjective probabilities. *Knowledge-Based Systems*, 83:170–175, July 2015.
- [41] Ronald R. Yager and Naif Alajlan. Some issues on the owa aggregation with importance weighted arguments. *Knowledge-Based Systems*, 100:89–96, May 2016.
- [42] R.R. Yager. On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making. *IEEE Transactions on Systems Man and Cybernetics*, 18(1):183–190, Jan 1988.
- [43] R.R. Yager. The power average operator. *Transactions on Systems, Man, and Cybernetics, Part A: Cybernetics*, 31:724–730, 2001.
- [44] R.R. Yager and D.P. Filev. Induced ordered weighted averaging operators. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, 29(2):141–150, Apr 1999.
- [45] L.A. Zadeh. A computational approach to fuzzy qualifiers in natural languages. *Computing and Mathematics with Applications*, 9(1):149–184, 1983.
- [46] Lotfi A. Zadeh. The concept of a linguistic variable and its applications to approximate reasoning - I. *Information Sciences*, 8:199–249, 1975.
- [47] Shang-Ming Zhou, Francisco Chiclana, Robert John, and Jonathan Garibaldi. Type-1 OWA operators for aggregating uncertain information with uncertain weights induced by type-2 linguistic quantifiers. *Fuzzy Sets and Systems*, 159(24):3281–3296, 2008.
- [48] Shang-Ming Zhou, Francisco Chiclana, Robert John, and Jonathan Garibaldi. Alpha-level aggregation: a practical approach to type-1 OWA operation for aggregating uncertain information with applications to breast cancer treatments. *IEEE Transactions on Knowledge and Data Engineering*, 10:1455–1468, 2011.