

Research Article

Opinion Mining from Online User Reviews Using Fuzzy Linguistic Hedges

Mita K. Dalal¹ and Mukesh A. Zaveri²

¹ Information Technology Department, Sarvajanic College of Engineering & Technology, Surat 395001, India

² Computer Engineering Department, S. V. National Institute of Technology, Surat 395007, India

Correspondence should be addressed to Mita K. Dalal; parikhmita@gmail.com

Received 29 August 2013; Accepted 1 January 2014; Published 20 February 2014

Academic Editor: Sebastian Ventura

Copyright © 2014 M. K. Dalal and M. A. Zaveri. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, there are several websites that allow customers to buy and post reviews of purchased products, which results in incremental accumulation of a lot of reviews written in natural language. Moreover, conversance with E-commerce and social media has raised the level of sophistication of online shoppers and it is common practice for them to compare competing brands of products before making a purchase. Prevailing factors such as availability of online reviews and raised end-user expectations have motivated the development of opinion mining systems that can automatically classify and summarize users' reviews. This paper proposes an opinion mining system that can be used for both binary and fine-grained sentiment classifications of user reviews. Feature-based sentiment classification is a multistep process that involves preprocessing to remove noise, extraction of features and corresponding descriptors, and tagging their polarity. The proposed technique extends the feature-based classification approach to incorporate the effect of various linguistic hedges by using fuzzy functions to emulate the effect of modifiers, concentrators, and dilators. Empirical studies indicate that the proposed system can perform reliable sentiment classification at various levels of granularity with high average accuracy of 89% for binary classification and 86% for fine-grained classification.

1. Introduction

In the present age, it has become common practice for people to communicate or express their opinions and feedbacks on various aspects affecting their daily life through some form of social media. An upsurge in online activities like blogging, social networking, emailing, review posting, and so forth has resulted in incremental accumulation of a lot of user-generated content. Most of these online interactions are in the form of natural language text. This in turn has led to increased research interest in content-organization and knowledge engineering tasks such as automatic classification, summarization, and opinion mining from web-based data.

Due to its high commercial importance, mining and summarizing of user reviews are a widely studied application [1–9]. The two main tasks involved in opinion mining regardless of the application are (1) identification of opinion-bearing phrases/sentences from free text and (2) tagging the sentiment polarity of opinionated phrases. The descriptors such as adjectives or adverbs describing the features present

in an opinion sentence mainly indicate the polarity of the expressed opinion. However, the strength and polarity of the opinionated phrases are also affected by the presence of linguistic hedges such as *modifiers* (e.g., “not”), *concentrators* (e.g., “very,” “extremely”), and *dilators* (e.g., “quite,” “almost,” and “nearly”). Zadeh developed the concept of fuzzy linguistic variables and linguistic hedges that modify the meaning and intensity of their operands [10, 11]. Recent papers in this field have also pointed out that the task of opinion mining is sensitive to such hedges and taking the effect of linguistic hedges into consideration can improve the efficiency of the sentiment classification task [8, 12–17].

In this paper, we have proposed an approach to perform fine-grained sentiment classification of online product reviews by incorporating the effect of fuzzy linguistic hedges on opinion descriptors. We have proposed novel fuzzy functions that emulate the effect of different linguistic hedges and incorporated them in the sentiment classification task.

Our opinion mining system involves various phases like (1) preprocessing phase, (2) feature-set generation phase,

and (3) fuzzy opinion classification phase based on fuzzy linguistic hedges. Empirical studies indicate that our sentiment mining approach can be successfully applied for binary as well as fine-grained sentiment classification of user reviews. In binary sentiment classification the reviews are classified into two output classes “*positive*” and “*negative*.” In fine-grained classification the reviews are classified into multiple output classes as “*very positive*,” “*positive*,” “*neutral*,” “*negative*,” and “*very negative*.” Our opinion mining system can also be used to generate comparative summaries of similar products [8]. Moreover the proposed fuzzy functions for emulating linguistic hedges give better accuracy than other contemporary approaches for hedge adjustment [13, 14, 16]. Note that while there are several recent papers on user review classification, there are relatively few which have explicitly proposed approaches to integrate the effect of linguistic hedges [13, 14, 16, 17].

The rest of the paper is organized as follows. Section 2 surveys related work done in the area of opinion mining, Section 3 describes the proposed framework for opinion mining using linguistic hedges, and Section 4 discusses the empirical evaluation of the proposed strategy and obtained results. Finally, we conclude and give directions for future work in this field.

2. Related Work

Mining opinions from user-generated reviews are a special application of natural language processing that requires automatic classification and summarization of text.

Automatic text classification is usually done by using a pre-labeled training set and applying various machine learning methods such as Naïve Bayesian [18], support vector machines [19], Artificial Neural Networks [20], or hybrid approaches [21, 22] that combine various machine learning methods to improve the efficiency of classification. Approaches to automatic text summarization have mainly focused on extracting significant sentences from text and can be broadly classified into four categories: (1) heuristics-based approaches that rely on a combination of heuristics such as cue/key/title/position [23–26], (2) semantics-based approaches such as lexical chains [27] and rhetorical parsing [28], (3) user query-oriented approaches typically used for retrieval in search engines and question-answering applications based on metrics such as maximal marginal relevance (MMR) [29, 30], and (4) cluster-oriented approaches that form clusters of sentences based on sentence similarity computations and then extract the *central* sentence of each cluster to include in the summary [31].

However the task of summarizing or classifying the sentiment reflected in users’ opinions is quite different from the text mining approaches mentioned above. It is not focused on generating extractive summaries or classifying entire documents based on topic-indicative words. Instead, sentiment mining involves such tasks as semantic feature-set generation, identifying opinion words (usually adjectives or adverbs) and associating them with corresponding features, determining the polarity of the feature-opinion pairs, and

finally aggregating the mined results to detect overall sentiment [1–9, 12, 32–34]. Users’ opinions are usually expressed informally in natural language and frequently contain errors in spelling and grammar. So, they require a lot of pre-processing to generate clean text [8, 12]. Moreover, feature extraction from user reviews requires language-dependent semantic processing such as parts-of-speech (POS) tagging [1, 5, 8, 12, 33] in addition to statistical frequency analysis. The POS tagging is usually done using a linguistic parser. For example, the link grammar parser [8, 35] and Stanford parser [12, 36] are well-known linguistic parsers. The nouns and noun phrases tagged by the parser become initial candidate features. Various approaches are used to extract the feature set useful for opinion mining. These approaches include frequent itemset identification using Apriori algorithm [1, 3, 5, 7, 32, 37], seed-set expansion using an initial seed set of features [12, 33], or multiwords-based [8, 38–41] frequent feature extraction.

Feature descriptors such as adjectives or adverbs are mainly indicative of the polarity of an opinion phrase. In [42], the authors proposed a method to determine orientation of adjectives based on conjunctions between adjectives. Statistical measures of word association such as pointwise mutual information (PMI) and latent semantic association (LSA) have also been used to determine semantic orientation [34, 43]. Another approach is to use a seed list of opinion words with previously known orientations [1, 12] and expand it based on lookup of synonyms and antonyms using some lexical resource like WordNet [44]. This approach is based on the observation that synonyms have similar orientations and antonyms have opposite orientations. It is important to note that the semantic orientation of a descriptor can differ depending on its usage. So, it is not enough to use corpus statistics to assign a fixed polarity to each descriptor. Hence, several recent papers have used the SentiWordNet tool [45] to determine opinion polarity [2, 8, 16, 33]. The advantage of using SentiWordNet is that it lists out all usages of a descriptor and assigns them a corresponding sentiment orientation triplet which indicates their positivity/objectivity/negativity scores.

In addition to opinion descriptors such as adjectives or adverbs, the orientation (i.e., polarity) and strength of an opinion phrase are sensitive to the presence of linguistic hedges [10, 14, 46]. Some authors refer to linguistic hedges as contextual valence shifters and have demonstrated that they can affect the valence (polarity) of a linguistic phrase [13, 14]. Moreover it has been shown that the accuracy of opinion classification can be improved by augmenting the simple positive/negative term counting approach to incorporate the effect of such hedges [13, 14]. In [16], the authors have used a hybrid scoring technique based on linear combination of PMI, SentiWordNet, and manually assigned scores to derive the initial sentiment value of an opinionated phrase, which is then adjusted using fuzzy functions when hedges are present. In this paper we have proposed alternative fuzzy functions to incorporate the effect of linguistic hedges. Our approach achieves higher accuracy compared to contemporary approaches.

3. Proposed Opinion Mining System

This section discusses the design of our proposed opinion mining system based on fuzzy linguistic hedges. Our opinion mining system automatically extracts opinionated phrases from unstructured user reviews and classifies them based on their sentiment. Moreover while assigning a sentiment score to a phrase, it takes into account the differences in intensity or polarity between opinionated phrases describing a feature “ f ” such as “ f is good” (no hedge), “ f is *extremely* good” (concentrating/intensifying hedge), “ f is *quite* good” (dilating/diminishing hedge), and “ f is *not* good” (modifying/inverting hedge).

We performed opinion mining on a dataset of online user reviews of various products which was collected using a web crawler. As depicted in Figure 1, our system consists of three major phases. These phases are (1) preprocessing phase, (2) feature generation phase, and (3) fuzzy opinion classification phase.

3.1. Preprocessing Phase. User-generated online reviews require preprocessing to remove noise [8, 12] before the mining process can be performed. This is because these reviews are usually short informal texts written by nonexperts and frequently contain mistakes in spelling, grammar, use of nondictionary words such as abbreviations or acronyms of common terms, mistakes in punctuation, incorrect capitalization, and so forth.

Since we need to perform POS tagging in the next phase using a linguistic parser we perform cleaning tasks such as spell-error correction using a standard word processor, sentence boundary detection [12], and repetitive punctuation conflation [8]. Syntactically correct sentences end with predefined punctuations such as full stop (.), interrogation mark (?), or exclamation mark (!). Sentence boundary detection involves various tasks such as identifying end of sentences on the basis of correct punctuations and disambiguating the full stop (.) from decimal points and abbreviated endings (e.g., “Prof.,” “Pvt.”). We also capitalize the first letter of each new sentence. Bloggers sometimes overuse a punctuation symbol for emphasis. In such cases, the repetitive symbol is conflated to a single occurrence [8]. For example, a review posted by a blogger may read as follows:

“the display is somewhat spotty!!!”

After preprocessing, the sentence would read as follows:

“The display is somewhat spotty!”

In the above sentence, the first letter has been capitalized and the repetitive exclamation mark (“!!!”) has been conflated so that it occurs only once (“!”).

Thus, preprocessing steps generate sentences which can be parsed automatically by the linguistic parser. Moreover, product reviews often quote abbreviations and acronyms relevant to the domain which cannot be found in standard language dictionaries. These cannot be considered as spelling mistakes. So, to make our system more fault tolerant we also generate a domain-specific resource of frequently occurring

TABLE 1: Partial list of acronyms for “smartphones.”

Acronym	Expanded form
iOS	iPhone operating system
HD video	High definition video
LTE	Long term evaluation
TFT	Thin film transistor
LCD	Liquid crystal display
AUX cable	Auxiliary cable
HDR	High dynamic range

abbreviations and acronyms [47] to augment the standard word dictionary. If a nondictionary word frequently occurs in the review dataset, it is examined by a human expert and added to the domain resource if found relevant. For example, Table 1 shows a partial list of acronyms that were extracted from our user review dataset for the “smartphones” product.

3.2. Feature-Set Generation Phase. In this phase, we generate the feature set for opinion mining from the cleaned review sentences generated in the preprocessing phase. Since the spelling and punctuation errors have been removed and sentence boundaries have been clearly identified, we now parse these sentences using the link grammar parser [35]. The parser outputs POS (parts-of-speech) tagged output. Frequently occurring nouns (N) and noun phrases (NP) are treated as features, while the adjective or adverb modifiers describing them are treated as opinion words or descriptors [8]. Additionally we also take into account any linguistic hedges preceding the descriptors. For example, consider the following review sentence for a “*smartphone*” product:

“The call quality is extremely good and navigation is comfortable but the body is somewhat fragile.”

When this review sentence is parsed using the link grammar parser, we get an output like

“The call [.n] quality [.n] is [.v] extremely good [.a] and navigation [.n] is [.v] comfortable [.a] but the body [.n] is [.v] somewhat fragile [.a].”

In the above sentence, [.n] indicates noun, [.a] indicates adjective, and [.v] indicates verb. Thus, “*call quality*” can be interpreted as a noun feature which is described by the adjective descriptor “*good*.” Similarly, “*navigation*” and “*body*” are features described by the descriptors “*comfortable*” and “*fragile*,” respectively. Moreover, the descriptor “*good*” is preceded by the concentrator hedge “*extremely*” and the descriptor “*fragile*” is preceded by the dilator hedge “*some-what*,” while the descriptor “*comfortable*” has no preceding hedge in this particular review sentence.

The mined feature set is tabulated in an FOLH table (feature orientation table with linguistic hedges). Table 2 shows the FOLH table entries for some of the most frequently commented upon features from the user review set for the “*smartphone*” product.

The FOLH table stores the product features as well as the descriptors and modifying hedges corresponding to

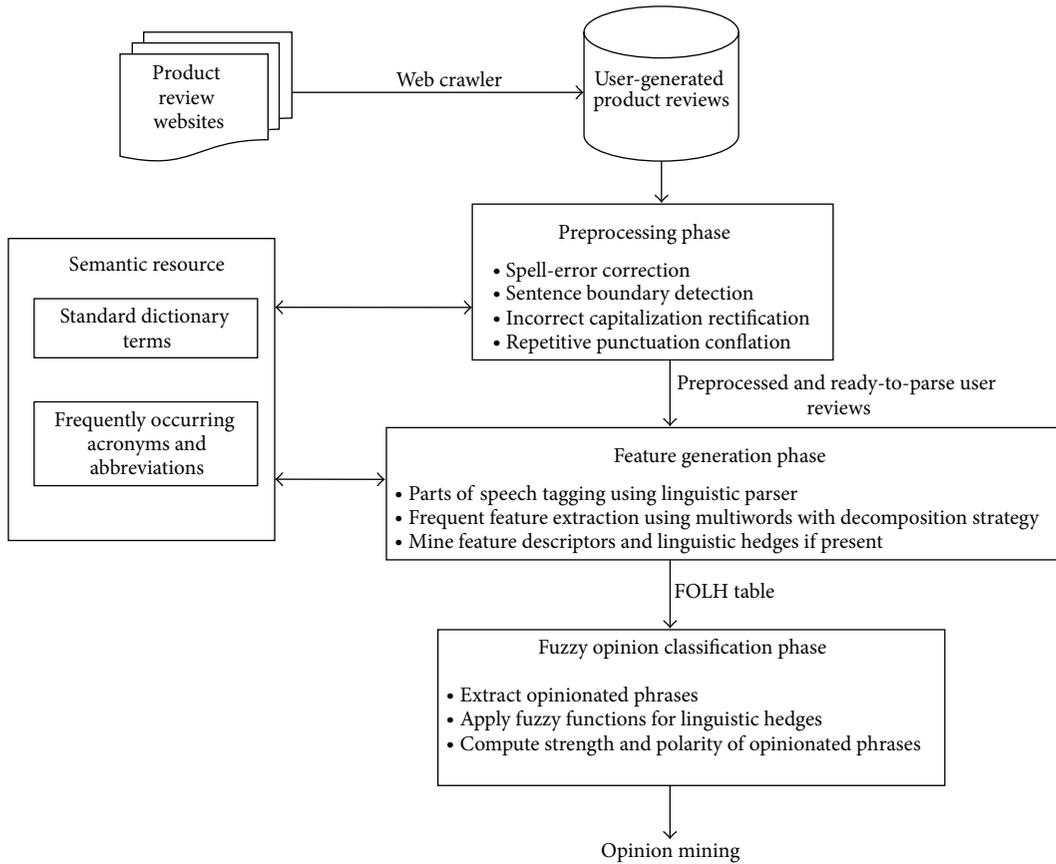


FIGURE 1: System design for opinion mining.

TABLE 2: Partial feature orientation table with linguistic hedges for smartphone products.

Feature	Descriptors with positive polarity ($P = 1$)	Descriptors with negative polarity ($P = -1$)	Linguistic hedges (if present)		
			Modifier (inverter)	Concentrator	Dilator
Call quality	Good, excellent, and satisfactory	Poor, bad	Not, never	Very, extremely	Quite, hardly
Body/design/build	Sleek, lightweight, thin, slim, beautiful, sturdy, striking, and gorgeous	Heavy, bulky, and fragile		Very, absolutely	Somewhat, quite, and almost
Screen/touchscreen/display/retina display	Nice, great, sensitive, awesome, clear, and bright	Dull, bad, and spotty	Not	Highly, incredibly	Quite
Camera/phone camera/digital camera	Awesome, good, superior, and high-resolution	Low resolution, inferior	Not	Very, positively	
User interface	Friendly, attractive, good, and lovely	Bad, poor	Not so	Highly, very	More or less
Navigation	Comfortable, intuitive, easy, and fast	Bad, difficult, slow, and jumpy		Very, significantly	Quite

the features mined from the training set of reviews. For example, consider the first entry in Table 2. It indicates that the linguistic variable “call quality” is a smartphone feature. This feature can take on fuzzy values like “good,” “excellent,” or “satisfactory” which have a positive polarity, or it can take

on fuzzy values like “poor” and “bad” which have a negative polarity. In addition, the intensity of the fuzzy values describing the feature can be increased by concentrator linguistic hedges such as “very” and “extremely” or decreased by dilator linguistic hedges “somewhat” and “hardly.” The sentiment

polarity can be reversed by the inverter hedges “not” and “never.”

Frequently occurring semantic word sequences are treated as multiword features [8, 38–41]. For example, in the previous example, “*call quality*” is a multiword feature. We use the multiword with decomposition strategy approach [39], for feature extraction, as it requires lesser pruning and improves classification accuracy compared to Apriori-based and seed-set expansion-based approaches [8]. The orientation and initial sentiment value of the corresponding descriptors are determined using the SentiWordNet tool [2, 8, 33, 45]. For example, the SentiWordNet score for adjective “*fragile*” (as used to describe *body* in the smartphone review sentence) is given by the triplet (P: 0, O: 0.375, and N: 0.625) which indicates its positive, objective, and negative score. Since the negative sentiment value is highest in the triplet, “*fragile*” is assigned a polarity of “-1” that indicates negative orientation and an initial sentiment intensity value “0.625” which is used in the next phase.

In the FOLH table, semantically similar features are clubbed together by human expert to avoid redundancy and to get a more accurate value of occurrence frequency [8]. It is important to note that several acronyms (e.g., *HD video*, *LCD screen*, etc.) identified during the preprocessing phase are actually multiword features and are thus added to the feature set. The FOLH table is used in the next phase to compute the overall sentiment score of a user review and classify it. Since hedges are generic terms which could be combined with any feature-descriptor pair, a consolidated list of hedges for each of the three categories (i.e., modifier, concentrator, and dilator) is prepared.

3.3. Fuzzy Opinion Classification Phase. In this phase we perform fine-grained classification of users’ reviews. The reviews are classified as *very positive*, *positive*, *neutral*, *negative*, or *very negative*. We classify a new user review based on its fuzzy sentiment score whose computation requires three steps: (1) extract features, associated descriptors, and hedges from the review based on FOLH table lookup, (2) identify the polarity and initial value of the feature descriptors based on SentiWordNet score, and (3) calculate overall sentiment score using fuzzy functions to incorporate the effect of linguistic hedges.

The first two steps are performed as explained in Section 3.2. As discussed earlier, we consider the SentiWordNet score of a feature descriptor as its initial fuzzy score $\mu(s)$. If the descriptor has a preceding hedge, its modified fuzzy score is calculated using

$$f(\mu(s)) = 1 - (1 - \mu(s))^\delta. \quad (1)$$

Similar to Zadeh’s proposition [10], if the hedge is a concentrator, we choose $\delta = 2$ which gives us modified fuzzy concentrator score as indicated in (2), while if the hedge is

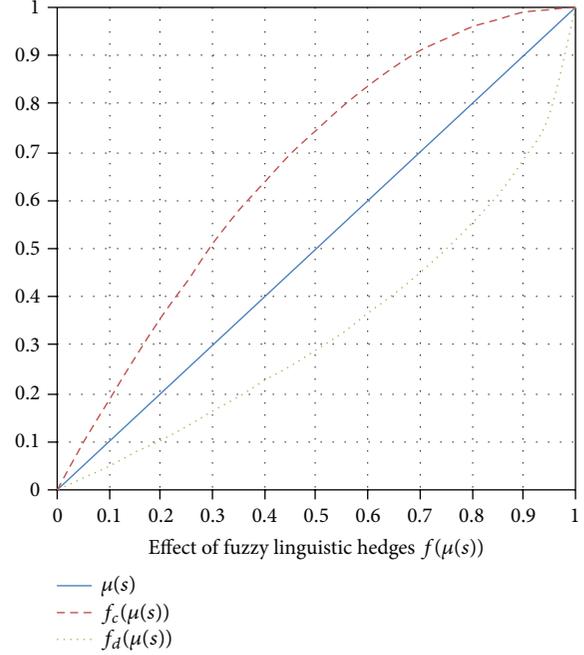


FIGURE 2: Depiction of linguistic hedges.

a dilator we choose $\delta = 1/2$ which gives us modified fuzzy dilator score as indicated in (3):

$$f_c(\mu(s)) = 1 - (1 - \mu(s))^2, \quad (2)$$

$$f_d(\mu(s)) = 1 - (1 - \mu(s))^{1/2}. \quad (3)$$

Let us revisit the smartphone review sentence “*The body is fragile.*” As explained in Section 3.2, the initial sentiment score for the descriptor “*fragile*” obtained using SentiWordNet is $\mu(s) = 0.625$. If this descriptor is preceded by a concentrator linguistic hedge, for example, “*very fragile*,” then its modified fuzzy score is obtained using (2) as $f_c(\mu(s)) = 0.8593$. Similarly, if this descriptor is preceded by a dilator linguistic hedge, for example, “*somewhat fragile*,” then its modified fuzzy score is obtained using (3) as $f_d(\mu(s)) = 0.3876$. Thus, the intensity level of a descriptor is adjusted on the basis of the linguistic hedge, whenever such hedges are present in a review sentence.

Figure 2 depicts the effect of applying fuzzy linguistic hedges such as concentrators and dilators as per (2) and (3), respectively.

The proposed fuzzy functions have several desirable properties as listed below.

Property 1. Consider $\forall x, y \in [0, 1]$ if $x < y$, $f_c(x) < f_c(y)$ and $f_d(x) < f_d(y)$.

Property 2. Consider $\forall x \in [0, 1]$ $f_d(x) < x < f_c(x)$.

Property 3. Consider $\forall x, f(x) \in [0, 1]$.

Let x and y indicate the initial sentiment values of a feature descriptor which are to be modified using the proposed functions for fuzzy linguistic hedges. From Property 1 it becomes clear that both the concentrator and dilator fuzzy functions are strictly increasing in the interval $[0, 1]$. Moreover, as indicated by Property 2, the dilator function decreases the value of the input sentiment variable while the concentrator function increases its value. Property 3 indicates that even after applying the fuzzy functions the output value $f(x)$ remains in the normalized range of $[0, 1]$.

Let U represent the complete feature set of a product. Suppose that a user review has comments on a subset F of the feature set. Further, let H represent the subset of F which is preceded by concentrator or dilator linguistic hedges, while N represents the subset of F not preceded by these hedges. Thus, $F \subset U$ and $F = H \cup N$.

Now, the average fuzzy sentiment score is calculated as shown in

$$\beta_{\text{avg}} = \frac{\sum_{i=1}^{|H|} p_i f(\mu_i(s)) + \sum_{k=1}^{|N|} p_k \mu_k(s)}{|F|}. \quad (4)$$

In (4), the first term of the numerator is derived from (1) and accounts for the descriptors which have been modified by hedges (concentrator or dilator as applicable), while the second term of the numerator accounts for the rest of the descriptors. The term “ p_i ” indicates the polarity of the i th feature descriptor which needs to be looked up from the FOLH table. If the polarity is positive, then its value is $+1$, and if the polarity is negative, its value is -1 . Note that (1)–(3) are only applicable to concentrator and dilator hedges. If there is an “inverter” hedge (e.g., “not”) preceding a feature descriptor, it is accounted for simply by reversing the value of polarity indicator “ p_i .” Thus an inverter hedge only changes the orientation of a sentiment phrase without affecting its magnitude.

The value of “ β_{avg} ” calculated using (4) falls in the range $[-1, 1]$. We further normalize this value using min-max normalization [48] to map it to the range $[0, 1]$. Upon applying min-max normalization to “ β_{avg} ,” we get the normalized fuzzy bias value “ β_N ” ($\beta_N \in [0, 1]$) as indicated in

$$\beta_N = \frac{\beta_{\text{avg}} + 1}{2}. \quad (5)$$

Once the value of β_N is computed, the opinion class C can be determined using the following rule set:

- if $\beta_N \geq 0$ and $\beta_N \leq 0.25$, then $C =$ “very negative,” else
- if $\beta_N > 0.25$ and $\beta_N < 0.5$, then $C =$ “negative,” else
- if $\beta_N = 0.5$, then $C =$ “neutral,” else
- if $\beta_N > 0.5$ and $\beta_N \leq 0.75$, then $C =$ “positive,” else
- if $\beta_N > 0.75$ and $\beta_N \leq 1$, then $C =$ “very positive.”

The accuracy of the sentiment classification task is verified by comparing the class assigned by our opinion miner with the star rating assigned by the user to that review. The next phase discusses the empirical evaluation of our proposed method.

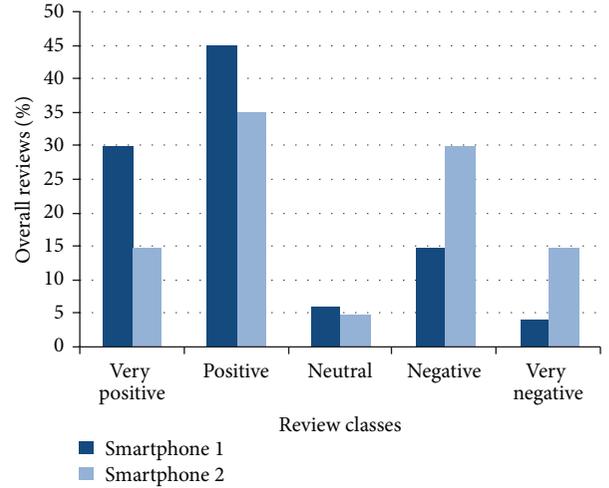


FIGURE 3: Fine-grained classification of user reviews.

4. Empirical Evaluation and Results

This section presents the results of empirical evaluation of our opinion mining strategy. In order to evaluate our approach, we used a dataset of over 3000 user-generated product reviews crawled from different websites. The review database consisted of user-generated reviews for four types of products (i.e., tablets, E-book readers, smartphones, and laptops) of different brands. We selected websites where, in addition to review text, the users also give a rating (1–5 stars) to their review. We use 30% of the review database as training set and 70% as the test set. As explained in Sections 3.1 and 3.2, we first preprocess the review text, extract product features, and generate the FOLH table using the training set of user product reviews. Then we perform classification on the test set of reviews using the equations and rule set derived in Section 3.3. The user-assigned 5-star rating is used as a basis to evaluate the accuracy of the proposed opinion mining system after the classification is complete. It is important to note that, unlike text classifiers based on supervised machine learning methods like Naïve Bayesian or SVM, the feature-based approach does not require a labeled training set for performing the classification.

Once the reliability of the opinion mining system is established, it can be used to automatically extract opinionated sentences from user reviews, perform fine-grained classification, and generate overall or feature-based comparative product summaries. For example, Figure 3 depicts the overall fine-grained sentiment classification-based comparative summary for two models of smartphones. It is clear from Figure 3 that “*smartphone 1*” is more popular among users since it has significantly more positive reviews compared to “*smartphone 2*.”

Figure 4 depicts the partial feature-based comparison of two smartphone products, based on some of the most frequently commented features wherein granularity of classification was reduced to improve readability. In this example, featurewise comparative product summary was generated by considering $\beta_N \geq 0.5$ as positive and $\beta_N < 0.5$ as negative.

TABLE 3: Accuracy of feature-based opinion classification using linguistic hedges.

Dataset	Approaches to feature-based classification using linguistic hedges					
	Approach 1: valence points adjustment approach		Approach 2: Vo and Ock's fuzzy adjustment functions		Approach 3: our proposed method	
	Binary classification accuracy (%)	Fine-grained classification accuracy (%)	Binary classification accuracy (%)	Fine-grained classification accuracy (%)	Binary classification accuracy (%)	Fine-grained classification accuracy (%)
Tablets	80.87	71.62	87.28	79.73	90.16	87.45
E-book readers	76.21	65.89	77.37	72.95	85.44	82.72
Smartphones	88.32	74.36	89.16	83.93	92.56	89.95
Laptops	84.26	73.32	86.67	79.65	88.67	85.77
Average	82.42	71.30	85.12	79.07	89.21	86.47

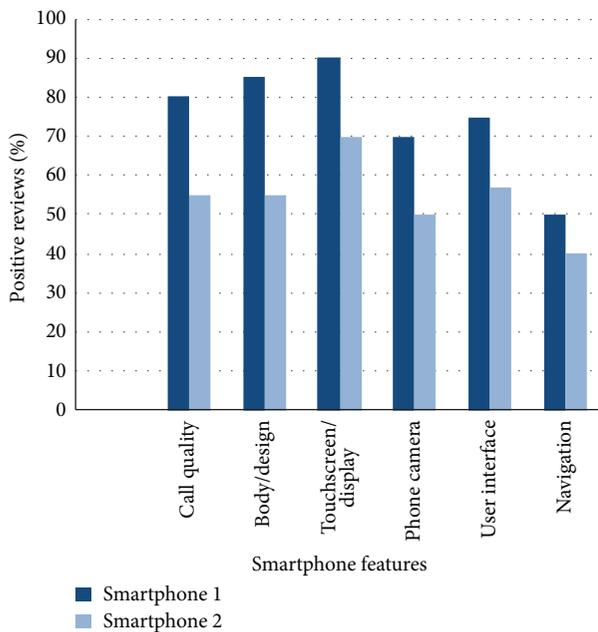


FIGURE 4: Comparative featurewise summary of two smartphones.

We evaluated the efficiency of our opinion mining system when used for binary as well as fine-grained sentiment classification. To evaluate the effectiveness of our approach for incorporating fuzzy linguistic hedges, we compared it with two other approaches: (1) valence points adjustment approach [13, 14] and (2) Vo and Ock's fuzzy adjustment approach [16].

The valence points adjustment approach is a simple hedge adjustment method that was proposed by Polanyi and Zaenen [13]. This method of valence adjustment has also been used by Kennedy and Inkpen in their system for rating movie reviews [14]. According to the valence points adjustment approach, all positive sentiment terms are given an initial or base value of 2 [13]. If this term is preceded by a concentrator (intensifier) in the same phrase its value becomes 3, while if it is preceded by a dilator (diminisher) its value becomes 1. Similarly, all negative sentiment terms are given a base value of -2 which are adjusted to values -1 and -3 if preceded by diminisher or intensifier hedges, respectively [13, 14].

Vo and Ock have proposed fuzzy adjustment functions [16] for incorporating the effect of hedges. They considered five categories of hedges (increase, decrease, invert, invert increase, and invert decrease) and have proposed fuzzy functions for each category [16].

The feature-based product review classification approach is augmented with the three hedge adjustment approaches and their classification accuracies are compared when applied to the task of binary as well as fine-grained sentiment classification of product reviews. The results of comparison of the three approaches are tabulated in Table 3.

It can be observed from Table 3 that all three approaches give acceptable accuracies (over 82%) when used for binary classification of user reviews where sentiment polarity is simply classified as "positive" or "negative." However, our proposed approach performs binary classification with higher average accuracy (89%) than the other two approaches.

When used for fine-grained classification, all three approaches tend to deteriorate in accuracy. This is understandable because increasing the number of categories for sentiment classification tends to result in more number of errors near the boundaries of adjacent classes (e.g., between "very negative" and "negative"). Here again, our proposed approach proves to be more robust than the other two approaches. Empirical results tabulated in Table 3 indicate that the accuracy of approach 1 decreased by approximately 11% while the accuracy of approach 2 decreased by approximately 6% over the test dataset when used for fine-grained classification wherein the number of output classes was increased. In contrast, our proposed approach shows only a 3% decline in accuracy. Our approach gives high accuracy of over 86% when used for fine-grained review sentiment classification and clearly outperforms the other two approaches. *Thus, the proposed opinion mining system successfully incorporates the effect of linguistic hedges and performs sentiment classification of reviews with acceptable accuracy.*

At present we have considered all online reviews to be of equal authenticity while performing the opinion mining. However, in future we would like to build an enhanced opinion mining system that calculates the weight of an opinion by establishing its authenticity. On some blogs, a user's initial or base review is often rated by other readers

by simply clicking on an *Agree/Thumbs Up* symbol to express agreement or a *Disagree/Thumbs Down* symbol to express disagreement. Sometimes the comments are further commented upon by other reviewers, thus forming chains of comments. Performing opinion mining on such chains can establish the *authenticity* of the initial review. For example, a sham review written by a competitor discrediting a rival's product would receive several "*Disagree*" comments by other readers. Spoof reviews can also jeopardize the recommendation system of an online shopping site. In future, we would like to enhance our opinion mining system to take into consideration such secondary comments to refine the weight of the base opinion, which in turn can be used to generate a reliable "recommendation system" for online shoppers.

5. Conclusion

Empirical results indicate that the proposed opinion mining system performs both binary and fine-grained sentiment classifications of user reviews with high accuracy. The proposed functions for emulating fuzzy linguistic hedges could be successfully incorporated into the sentiment classification task. Moreover, our approach significantly outperforms other contemporary approaches especially when the granularity of the sentiment classification task is increased.

In future, we would like to build an advanced opinion mining system capable of rating the authenticity of a user review based on mining opinion threads of secondary reviewers.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04)*, pp. 168–177, August 2004.
- [2] M. A. Jahiruddin, M. N. Doja, and T. Ahmad, "Feature and opinion mining for customer review summarization," in *Proceedings of the 3rd International Conference on Pattern Recognition and Machine Intelligence (PReMI '09)*, vol. 5909 of *Lecture Notes in Computer Science*, pp. 219–224, 2009.
- [3] S. Shi and Y. Wang, "A product features mining method based on association rules and the degree of property co-occurrence," in *Proceedings of the International Conference on Computer Science and Network Technology (ICCSNT '11)*, vol. 2, pp. 1190–1194, December 2011.
- [4] S. Huang, X. Liu, X. Peng, and Z. Niu, "Fine-grained product features extraction and categorization in reviews opinion mining," in *Proceedings of the 12th IEEE International Conference on Data Mining Workshops (ICDMW '12)*, pp. 680–686, 2012.
- [5] C.-P. Wei, Y.-M. Chen, C.-S. Yang, and C. C. Yang, "Understanding what concerns consumers: a semantic approach to product feature extraction from consumer reviews," *Information Systems and e-Business Management*, vol. 8, no. 2, pp. 149–167, 2010.
- [6] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP '05)*, pp. 339–346, October 2005.
- [7] W. Y. Kim, J. S. Ryu, K. I. Kim, and U. M. Kim, "A method for opinion mining of product reviews using association rules," in *ACM International Conference on Interaction Sciences: Information Technology, Culture and Human (ICIS '09)*, pp. 270–274, November 2009.
- [8] M. K. Dalal and M. A. Zaveri, "Semisupervised learning based opinion summarization and classification for online product reviews," *Applied Computational Intelligence and Soft Computing*, vol. 2013, Article ID 910706, 8 pages, 2013.
- [9] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," in *Proceedings of the International Conference on Empirical Methods in Natural Language Processing (EMNLP '02)*, pp. 79–86, 2002.
- [10] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-II," *Information Sciences*, vol. 8, no. 4, part 3, pp. 301–357, 1975.
- [11] V. N. Huynh, T. B. Ho, and Y. Nakamori, "A parametric representation of linguistic hedges in Zadeh's fuzzy logic," *International Journal of Approximate Reasoning*, vol. 30, no. 3, pp. 203–223, 2002.
- [12] L. Dey and Sk. M. Haque, "Opinion mining from noisy text data," *International Journal on Document Analysis and Recognition*, vol. 12, no. 3, pp. 205–226, 2009.
- [13] L. Polanyi and A. Zaenen, "Contextual valence shifters," in *Computing Attitude and Affect in Text: Theory and Applications*, vol. 20 of *The Information Retrieval Series*, pp. 1–10, 2006.
- [14] A. Kennedy and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters," *Computational Intelligence*, vol. 22, no. 2, pp. 110–125, 2006.
- [15] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.
- [16] A.-D. Vo and C.-Y. Ock, "Sentiment classification: a combination of PMI, sentiWordNet and fuzzy function," in *Proceedings of the 4th International Conference on Computational Collective Intelligence Technologies and Applications (ICCCI '12)*, vol. 7654, part 2 of *Lecture Notes in Computer Science*, pp. 373–382, 2012.
- [17] S. Nadali, M. A. A. Murad, and R. A. Kadir, "Sentiment classification of customer reviews based on fuzzy logic," in *Proceedings of the International Symposium on Information Technology (ITSim'10)*, pp. 1037–1044, mys, June 2010.
- [18] S. Kim, K. Han, H. Rim, and S. H. Myaeng, "Some effective techniques for naïve bayes text classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 11, pp. 1457–1466, 2006.
- [19] Z.-Q. Wang, X. Sun, D.-X. Zhang, and X. Li, "An optimal SVM-based text classification algorithm," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 1378–1381, August 2006.
- [20] Z. Wang, Y. He, and M. Jiang, "A comparison among three neural networks for text classification," in *Proceedings of the 8th International Conference on Signal Processing (ICSP '06)*, pp. 1883–1886, November 2006.
- [21] D. Isa, L. H. Lee, V. P. Kallimani, and R. Rajkumar, "Text document preprocessing with the bayes formula for classification using the support vector machine," *IEEE Transactions on*

- Knowledge and Data Engineering*, vol. 20, no. 9, pp. 1264–1272, 2008.
- [22] R. D. Goyal, “Knowledge based neural network for text classification,” in *Proceedings of the IEEE International Conference on Granular Computing (GrC '07)*, pp. 542–547, November 2007.
- [23] H. P. Luhn, “The automatic creation of literature abstracts,” *IBM Journal of Research and Development*, vol. 2, pp. 159–165, 1958.
- [24] H. P. Edmundson, “New methods in automatic extracting,” *Journal of the ACM*, vol. 16, no. 2, pp. 264–285, 1969.
- [25] C.-Y. Lin and E. H. Hovy, “Manual and automatic evaluation of summaries,” in *Proceedings of the ACL-02 Workshop on Automatic Summarization*, vol. 4, pp. 45–51, 2002.
- [26] C.-Y. Lin and E. H. Hovy, “Identifying topics by position,” in *Proceedings of the 5th Conference on Applied Natural Language Processing*, pp. 283–290, 1997.
- [27] R. Barzilay and M. Elhadad, in *Proceedings of the ACL Workshop on Intelligent Scalable Text Summarization*, Using lexical chains for text summarization, Ed., pp. 10–17, 1997.
- [28] D. Marcu, “Improving summarization through rhetorical parsing tuning,” in *Proceedings of the 6th Workshop on Very Large Corpora*, pp. 206–215, 1998.
- [29] J. Carbonell and J. Goldstein, “The use of MMR, diversity-based reranking for reordering documents and producing summaries,” in *Proceedings of the ACM 21st International Conference on Research and Development in Information Retrieval*, pp. 335–336, 1998.
- [30] J. Lin, N. Madnani, and B. J. Dorr, “Putting the user in the loop: interactive maximal marginal relevance for query-focused summarization,” in *Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT'10)*, pp. 305–308, June 2010.
- [31] P.-Y. Zhang and C.-H. Li, “Automatic text summarization based on sentences clustering and extraction,” in *Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT '09)*, pp. 167–170, August 2009.
- [32] M. Hu and B. Liu, “Mining opinion features in customer reviews,” in *Proceedings of the 19th National Conference on Artificial Intelligence (IAAI '04)*, pp. 755–760, San Jose, Calif, USA.
- [33] L. Zhao and C. Li, “Ontology based opinion mining for movie reviews,” in *Proceedings of the 3rd International Conference on Knowledge Science, Engineering and Management*, pp. 204–214, 2009.
- [34] P. D. Turney, “Thumbs Up or Thumbs Down?: semantic orientation applied to unsupervised classification of reviews,” in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 417–424, 2002.
- [35] D. Sleator and D. Temperley, “Parsing english with a link grammar,” in *Proceedings of the 3rd International Workshop on Parsing Technologies*, pp. 1–14, 1993.
- [36] D. Klein and C. D. Manning, “Fast exact inference with a factored model for natural language parsing,” in *Advances in Neural Information Processing Systems (NIPS '02)*, pp. 3–10, MIT Press, Cambridge, Mass, USA, 2003.
- [37] R. Agrawal and R. Srikant, “Fast algorithm for mining association rules,” in *Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487–499, 1994.
- [38] K. W. Church and P. Hanks, “Word association norms, mutual information and lexicography,” *Computational Linguistics*, vol. 16, no. 1, pp. 22–29, 1990.
- [39] W. Zhang, T. Yoshida, and X. Tang, “Text classification using multi-word features,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC '07)*, pp. 3519–3524, October 2007.
- [40] M. K. Dalal and M. A. Zaveri, “Automatic text classification of sports blog data,” in *Proceedings of the Computing, Communications and Applications Conference (ComComAp '12)*, pp. 219–222, January 2012.
- [41] W. Zhang, T. Yoshida, and X. Tang, “TFIDE, LSI and multi-word in information retrieval and text categorization,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '08)*, pp. 108–113, October 2008.
- [42] V. Hatzivassiloglou and K. Mckeown, “Predicting the semantic orientation of adjectives,” in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics (ACL '98)*, pp. 174–181, 1998.
- [43] P. D. Turney and M. L. Littman, “Measuring praise and criticism: inference of semantic orientation from association,” *ACM Transactions on Information Systems*, vol. 21, no. 4, pp. 315–346, 2003.
- [44] G. A. Miller, “WordNet: a lexical database for English,” *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [45] S. Baccianella, A. Esuli, and F. Sebastiani, “SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining,” in *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC '10)*, pp. 2200–2204, 2010.
- [46] T. Zamali, M. A. Lazim, and M. T. A. Osman, “Sensitivity analysis using fuzzy linguistic hedges,” in *Proceedings of the IEEE Symposium on Humanities, Science and Engineering Research*, pp. 669–672, 2012.
- [47] M. K. Dalal and M. A. Zaveri, “Automatic classification of unstructured blog text,” in *Journal of Intelligent Learning Systems and Applications*, vol. 5, no. 2, pp. 108–114, 2013.
- [48] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, chapter 2, Elsevier, 2nd edition, 2006.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

