Opinion Mining in Conversational Content within Web Discussions and Commentaries

Kristína Machová, Lukáš Marhefka

Opinion Mining in Conversational Content within Web Discussions and Commentaries

Kristína Machová[1] and Lukáš Marhefka[1],

[1] Dept. of Cybernetics and Artificial Intelligence, Technical University, Letná 9,
042 00, Košice, Slovakia,
kristina.machova@tuke.sk, lukas.marhefka@student.tuke.sk

Abstract. The paper focuses on the problem of opinion classification related to web discussions and commentaries. It introduces various approaches known in this field. It also describes novelty methods, which have been designed for short conversational content processing with emphasis on dynamic analysis. This dynamic analysis is focused mainly on processing of negations and intensifiers within the opinion analysis. The contribution describes implementations of these methods for the Slovak language. The Slovak dictionaries have been created to support these implementations of lexicon based approach. In addition, the paper describes test results of the presented implementations and discussion of these results as well.

Keywords: Opinion mining, conversational content, opinion classification, dynamic coefficient, n-grams.

1 Introduction

The social web or web with various forms of conversational content increases interactions between users. These interactions are provided in the form of "point-to-point" or "multicast" on-line web services, as chats, discussion forums, IRC (Internet Relay Chat), blog and micro-blog platforms and so on. These services create a big amount of text and therefore they offer interesting possibilities for research in the field of opinion classification.

Where are some interactions, there people influence each other. This influence concerns decision making about various life situations, for example decisions about purchase or production and selling some products, decisions about voting political representatives and so on. This decision making can be supported by information mining from web discussions, reviews, conversations about matter of our interest. Thus, our decision processes can be encouraged by web conversational content. The mentioned conversational content can provide us also with cultural information about films, books and cultural events. On the other hand, it can be a source of some information connected with safety issues, for example suspicious activities or characters, connected with racism, pedophilia or terrorism.

But sometimes these web discussions can be too long and some parts of these discussions can be less informative. That is why there is a need of some tools for

automatic analysis of web discussions from the point of positivity or negativity. Our implementation of various methods of opinion analysis provides one such tool.

An opinion represents some positive or negative attitude, view, approach, or emotion of some person – an opinion owner. Opinions are related to a given entity or some of its parts. The entity is some product, person, event, organization, etc. It is the object, we are talking about. It is composed from components. The entity can be some product, for example a mobile. It has some parts – attributes, for example display, design and size, functionalities and so on.

There can be identified a set of different tasks within the field of the opinion analysis:

- *emotion detection* (Does the conversation obtain some emotion?)
- *opinion spam detection* (Is the content of some discussion contribution informative?)
- *subjectivity analysis* (Does the conversation obtain subjective opinion?)
- *opinion polarity classification* (Is the conversation about given subject positive or negative?)
- *theme modeling* (What is the theme/object of this conversation? Is this theme a suspicious one?)
- *authorship identification* (Who is the opinion author? What kind of person is the author of a given contribution?)

Our approach, presented within this paper, is focused on the problem of opinion polarity classification, shortly on opinion classification.

2 Related Works

There are some approaches similar to our approach to opinion classification [2], [11], and [12]. New technical contribution of our approach in comparison with existing works will be discussed within this section. The most similar to our approach is Taboada at al: "Lexicon-Based Methods for Sentiment Analysis" in [11]. It uses a dictionary of words annotated with their orientation (polarity). This approach splits the dictionary into more sub-dictionaries according to word classes (adjectives, nouns, verbs and adverbs). We use only one dictionary (for dynamic coefficient) or two dictionaries (for n-grams). The novelty of our approach is, that the dictionary is generated directly from web discussions, which increases the precision of opinion analysis of the given discussions. In [11], intensification is provided by increasing (respectively decreasing) the semantic intensity of *neighboring* lexical items using a special dictionary of intensifiers. Our approach is more flexible because an intensifier and the related word need not to be neighbors. They can take any position within one lexical unit, while their distance is limited by the given dynamic coefficient. The intensifier can be located before or after the related word. Within our approach, not only intensification, but also negation processing is different, based on processing various combinations of words (lexical units) defined with the aid of dynamic coefficient. All lexical units are sequentially classified into six categories (3, 2, 1, -1, -

2, -3). Three of them represents positive polarity and the other three categories negative polarity (with meaning: strong + intensifier, strong or gentle).

Another approach presented in Thelwall at al: "Sentiment strength detection in short informal text" [12] is focused on the SentiStrength detection algorithm, which solves some problems connected with sentiment analysis (generation of the sentiment strength list, optimization of the sentiment word strengths, allocation of the miss words, spelling correction, creation of the booster word list, the negating word list and emoticon list, repeated letters processing and ignoring negative emotion in questions). They devote more effort to correction of non-standard spelling in informal texts. Our approach is not aiming at correct spelling of contributions since the dictionary can easily accommodate some misspelled words as well. Within our approach, the dictionary is generated directly from the analysed discussion and therefore words with sentiment are accepted in spite of common mistakes (for example very common typo using "y" instead of "i" and vice versa in the Slovak language). The algorithm described in [12] was tested on data from MySpace and on wide variety of themes. We have provided our tests on data from narrow domains and therefore our results were a little bit better.

Paper "Learning with compositional semantics as structural inference for substantial sentiment analysis" [2] written by Choi and Cardie is focused on sentiment analysis similarly to our approach. It presents a novel learning based approach that incorporates inference rules inspired by compositional semantics into the learning procedure. Our approach differs from this work because our method is dictionary based (we use simple bag-of-word approach) and it is not machine learning oriented. Our method incorporates surrounding of processed word up to distance K (maximally K neighboring words from the given word). Design of the method in [2] represents meaning of a composed expression as a function of the meanings of its parts within the compositional semantics. In our approach, these parts are lexical units, lengths and number of which are defined by the dynamic coefficient. The approach presented in [2] processes negations and intensification separately. This processing is made over the whole text with the aid of "voting-based inference". In our approach, the negation and intensifications are processed by the same mechanism of using dynamic coefficient. This processing is made separately in each lexical unit, not as the majority vote. Our approach does not use natural language processing. It is based on some statistical principles and so it better processes longer texts, which contain more than one sentence and very short texts (one sentence) can be analyzed with lower precision.

The machine learning approaches (Naïve Bayes classifier, Support Vector Machines (SVM) and Maximal Entropy) were used in [4], which is focused on the automatic classification of opinions from the micro-blog service Twitter. Within [4] the input data for SVM are represented by vectors with dimension **m**. Each item of this vector represents an attribute in the tweet. They used unigram extractor and so each attribute is one word. If this word is presented in the tweet, then its value in the vector is 1. Otherwise its value is 0. This substitution of the frequency of the given word occurrence by simple presence function leads to speedup of a tweet processing.

Sometimes, the introduced opinion analysis is denoted as opinion mining, because it focuses on the extraction of positive or negative attitude of a participant to the commented objects with the aid of mining techniques applied to text documents.

Opinion mining can be extended from the level of whole texts perception to the level of extraction of properties of those objects which match users' interests [3]. Parallel approach to opinion mining is sentiment analysis [10]. Deeper view on sentiment analysis, which is presented in [7], focuses on feature selection. Different approach to web discussion processing is represented by the estimation of authority degree of some information sources, for example of actors contributing to discussion forums or social networkss. An important technique for authoritative actors searching is visualization approach, which is introduced in [5]. Some effort was spent on semantically enriching algorithms for analysis of web discussion contributions by authors of [8]. Also dedicated information can be used as an interface to newsgroup discussions [9].

Nowadays, opinion analysis has become an important part of social networks analysis. Existing opinion analysis systems use large vocabularies for opinion classification into positive or negative answer categories. Such approach was used in [1]. Authors studied accuracy of the opinion analysis of Spanish documents originated in the field of economic. This approach uses a regression model for classification into negative or positive opinions. Authors studied how quality depends on the granularity of opinions and rules, which were used in the regression model. Another study [6] was focused on the possibility of using lesser granularity without any significant precision decrease. The results of this study show no substantial difference between one and two parameter regression models as well as no statistically significant difference between models with different granularity. Thus, for example, simpler models can be used with the sentiment scale reduced to five degrees only. The approach, presented in this paper, uses a scale with five degrees for opinion classification as well, but it differs from the previous approaches in vocabulary cardinality. Our work focuses on creating vocabularies with strong orientation on the discussion domain, not so large but created directly from live discussions. We do not use regression models.

3   Opinion Mining

The process of the opinion classification consists of minimally two subtasks:
1. *input text preprocessing* or transformation of discussion texts into lexical units which can be easily processed,
2. *opinion classification*, which represents determination (calculation) of the polarity (positive or negative) of a subjective opinion.

3.1   Input Text Preprocessing

There are several approaches how to parse an analyzed input text into smaller lexical units. The most used approaches are n-grams and part-of-speech tagging.

N-gram can be defined as a series of items from some sequence. From the semantic point of view, it can be a sequence of phones, characters or words. In practice, n-gram as a sequence of words is the most common. The sequence of two (three) words is called bigram (trigram). For the case, when more than three words (exactly *n*) are in the same sequence, such sequence is called n-gram. N-grams are

used in the wide scale of fields, as theoretical mathematics, biology, cartography, even in the field of music.

In the field of natural language processing, n-grams can be used for words prediction. The words prediction uses so called "n-grams model". This n-gram model calculates the probability of occurrence of the last word in an n-gram from the previous n-grams. Another way of using n-grams is plagiarism discovery by text dividing into smaller fragments. These fragments are represented by n-grams. These n-grams can be easily compared and consequently, the measure of similarity of comparing documents can be calculated. N-grams are often used for text categorization and also for effective searching for correct candidates of misspelled words. Our approach, presented within this contribution uses n-grams for splitting the web discussion contributions into lexical units.

The part-of-speech (POS) tagging represents the recognition of word class within the text on the basis of the given language attributes and relationships between word classes. The POS tagger is a program, which is able to read text and to assign word class to each word from the given text. Such word class assigning can be based on the word definition or the word position within the sentence. There are three types of POS taggers: rule taggers, stochastic taggers and transformation taggers. The rule tagging algorithms are based on sets of rules, which are used for tagging of the processed text. This tagged text can be used as a training set for stochastic approach. The stochastic taggers are based on probability of the given tag occurrence within the given text. The stochastic approach requires an input training set. The last one, transformation taggers represent combination of two previous approaches.

3.2 Opinion Classification

There are two basic approaches to opinion classification [11]:

- *machine learning approach* (classification based approach), which uses the statistical and machine learning methods
- *lexicon based approach*, which can be based on the dictionary (*dictionary based approach*) or corpus.

Machine Learning Approach. This approach is based on using some of well-known methods of machine learning, for example Naïve Bayes classifier, Support Vector Machines (SVM) and Maximal Entropy and so on.

Naïve Bayes is a simple classifier based on Bayes formula of conditional probability with the big degree of freedom. A problem is defined as a classification of a given tweet **d** to a class **c**. This class **c** represents positive or negative polarity of some contribution. The original Bayes theorem in the formula (1)

$$P(c \mid d) = \frac{P(c)P(d \mid c)}{P(d)} \qquad (1)$$

was modified taking into account the following fact: the probability of a tweet occurrence is 100% and so probability p(d|c) was replaced by probability of feature $\mathbf{f_i}$ in the class $\mathbf{c}$. The final Bayes formula for conversational content is following (2)

$$P_{NB}(c \mid d) = \frac{(P(c)\Sigma_{i=1}^{m}P(f_i \mid c)^{n_i(d)})}{P(d)}.$$

(2)

Maximal Entropy. Within the information theory, the entropy represents a measure of indeterminism. The principle of maximal entropy is an axiom of the Bayes probability theory. It states that the probability distribution, which can represent current state of knowledge in the best way, is the distribution with the highest information entropy. This model is described by the following formula (3)

$$P_{ME}(c \mid d, \lambda) = \frac{\exp(\Sigma_i \lambda_{i,c} F_{i,c}(d,c))}{\Sigma_{c'} \exp(\Sigma_i \lambda_{i,c} F_{i,c}(d,c))}.$$

(3)

Within this formula: $\mathbf{c}$ is positive (negative) class, $\mathbf{d}$ is a tweet and λ is a weighting vector. The weighting vector represents an importance of the given feature $\mathbf{F}$ for the given class assignment (usually the word with strong positive meaning). The weighting vector is set on the base of numerical optimization of its components by the maximization of conditional probabilities. The classification was provided by Standford classifier [13].

Support Vector Machines (SVM) is the groups of machine learning classification methods. The difference between the SVM and Perceptron is that SVM can be also used for nonlinear classification.

Dictionary Based Approach. This approach looks on the input text as a set of words. It does not take into account relationships between words within sentences or grammatical rules. Not every word has the same importance in the opinion classification. It is important to find within the processed text mainly words, which can express opinion in the best way. Such words are stored in the classification dictionary. The dictionary has to obtain some words with sentiment before starting the process of opinion classification. On the base of known polarity of such characteristic words, polarity of whole texts from conversation and finally polarity of the whole web discussion can be determined.

The classification dictionary is a database of words. These words are descriptors with marked influence on polarity determination, mainly adjectives, adverbs, nouns and verbs. The simplest dictionaries are the dictionaries which enable binary classification. More suitable are fuzzy dictionaries, which are able to determine not only polarity but also the strength of the polarity. Such dictionary can be created for only a given application or some known lexicons can be used, for example Word Net, WordNet-Affect, SenticNet, SentiWordNet and so on.

Each dictionary should contain words, which are common within web conversation forums, blogs, commentaries and so on including slang words, words without

diacritics and words with more common grammatical mistakes. Our approach, presented within this contribution is the dictionary based approach.

4 Negation and Intensification

Besides basic problems of opinion analysis as word subjectivity identification, word polarity (orientation) determination and determination of intensity of the polarity, each application of opinion classification has to solve the processing of negation and intensification. The basic problems can be simply solved with the aid of classification dictionaries. These dictionaries focus on those words, which are able to express subjectivity very well - mainly adjectives (e.g. 'extraordinary') and adverbs (e.g. 'awfully') are considered. On the other hand, other word classes must be considered as well in order to achieve satisfactory precision, for example nouns (e.g. 'crash') or verbs (e.g. 'damage'). The words with subjectivity are important for opinion classification. Therefore, they are identified and inserted into one vocabulary or into a set of vocabularies – one vocabulary for each word class (adjectives dictionary, adverbs dictionary, nouns dictionary and verbs dictionary). Words with subjectivity are inserted into the corresponding vocabulary together with their degree of polarity. The majority of opinion classification applications work with 5 to 10 degrees of polarity.

4.1 Negation Processing

There are many approaches to negation processing:
- Switch negation,
- Shift negation,
- Dynamic coefficient usage.

The *switch negation* is simply reversion of the polarity of the lexical item. The reversion represents changing the number sign of the polarity degree (from minus to plus and vice versa). There are many various words related to negation that need to be taken into account: *not, none, never, nothing*, which usually are situated next to a related word – item. But also other negations should be taken into account as *without, don't, lack* and so on, which can be situated in significant distance from the lexical item. These negations can be hardly processed by switch negation. What is more, switch negation can be insufficiently precise, because negation of a strong positive word rarely is a strong negative word and vice versa. More often the negation of a strong positive word is a slightly negative word and vice versa.

The *shift negation* focuses on the case, when negation of a strong positive (negative) word is not strong negative (positive). Instead of changing the sign, "shift negation" shifts polarity degree toward the opposite polarity by a fixed value. For example, if the shift value is "6", then negation can be calculated in the following way: "It is not completely dysfunctional ($-5 + 6 = +1$) but also not very useful ($+5 - 6 = -1$)."

The *dynamic coefficient* usage does not need to suppose that all negation words are situated next to related words – items. In this method, the text of conversation is divided into lexical units of the length given by dynamic coefficient. It does not matter how far the negation is from the related word within the same lexical unit. One sentence can be transformed into one or more lexical units.

4.2 Intensification

There are two different approaches to intensification:
- Dictionary based intensification,
- Intensification based on dynamic coefficient.

The *dictionary based intensification* supposes the dictionary of intensifications – words, which are able to increase (or decrease) the intensity of polarity. Each of intensifications should be stored in dictionary together with sign and number. This number can represent a percentage of changing of polarity intensity and sign represents the type of this change. This approach supposes that a lexical item is situated next to intensifier changing the polarity degree. This condition is fulfilled only in small number of cases.

The *intensification based on dynamic coefficient* does not suppose that a lexical item is situated next to intensifier changing the polarity degree. This kind of intensification does not depend on how far the negation is from the related word within the same lexical unit.

Our implementation uses dynamic coefficient for negation processing and also for intensification.

5 Dynamic Coefficient in Opinion Classification

First, a simple version of Opinion Classification Application (OCA) has been designed and implemented. This OCA represented the dictionary based approach to opinion classification with "*static coefficient*", which was defined by user as the parameter of the application. This static coefficient designates the length of a lexical unit (number of words from the text, which belong to one lexical unit) for processing within texts - contributions of discussion forums. This processing represents the classification to positive or negative opinions. The OCA solves basic problems of opinion classification (words with subjectivity identification, words polarity determination and polarity intensity determination) but also problems connected with word polarity reversion by negation and intensification of words. The solving of these problems desire to process a combination of words instead of separated words. A length of processed word combination can be represented by the static coefficient, but also by a dynamic coefficient.

The value of the "*dynamic coefficient*" K is being dynamically changed during processing of different lexical units. The dynamic coefficient adapts itself to the length of a lexical unit (sequence of words) under investigation. The value K represents the

number of words, which are included into the same word combination. In the case, when the value is higher than the number of words in the sentence, this value is dynamically decreased in order to ensure, that the combination contains only words from the investigated sentence, not from the beginning of the next sentence.

Three ways of the dynamic coefficient setting have been proposed. Before the process of opinion classification can start, user has to select one of the methods for coefficient determination, as it is illustrated in Fig 1.

Fig. 1. Three ways of the dynamic coefficient determination.

The first used way of dynamic coefficient setting is the calculation of the *average length of all sentences* of the discussion contribution text, which is analyzed. So, each contribution text from a live discussion can have the different value of the dynamic coefficient. The calculated value of the dynamic coefficient is used for processing of all sentences of the given contribution. Thus, each sentence of the same contribution is processed using the same value of the dynamic coefficient, although actual length of these sentences is different.

The second used way of dynamic coefficient setting is the calculation of the *half length of each sentence* of the discussion contribution text. If needed, the calculated value is rounded up. Each analyzed sentence is processed using a different value of the dynamic coefficient.

The last used way is the *hybrid approach*, which determines the value of the dynamic coefficient as an average of two values obtained from *Average length of all sentences* method and *Half length of each sentence* method, respectively.

6 N-grams Application to Opinion Classification

Our n-grams application to opinion classification belongs to dictionary approach. This application is oriented on the Slovak language, so it uses a dictionary of Slovak language words. The structure of this dictionary is given by four more important word classes for opinion analysis: adjectives, adverbs, nouns and verbs. The dictionary consists of two different parts. The first part contains adjectives, nouns and verbs. The second one contains adverbs and negations. The first part of the dictionary is used for solving basic problems of opinion classification. The second part of the dictionary is used for negation processing and intensification because the adverbs have in a language the function to increase (*"surprisingly nice"*) or to decrease (*"extremely low-class"*) intensity of a word polarity. The dictionary contains also some emoticons, which naturally can express emotions and opinions very well. Sometimes, the text analysis can be less clear and so emoticons can increase the precision of the classification. These emoticons are stored within the first part of the dictionary.

All words and emoticons from this first dictionary are quantified to polarity degree within the interval from -3 to 3. With respect to the second part of dictionary, intensifiers (adverbs) are assigned by value from -0.5 to 1 and negations are represented by value -2. The analyzed text is processed by the following way. All words from a text are compared with all words in the first dictionary, and in the case of match, the value of the word found in the dictionary is added into the overall sum of values of analyzed words. This sum is consequently multiplied by the second sum of values obtained from the second dictionary (for intensifiers and negations). This processing can be represented by formula (4)

$$P = \sum v(w_i^1)[1+\sum v(w_j^2)] \; . \tag{4}$$

Within this formula:
P   …is the polarity degree of analyzed text
$v(w_i^1)$ …is the value of word w_i of text found in the first part of the dictionary
$v(w_j^2)$ …is the value of word w_j of text found in the second part of the dictionary.

For example:
- the lexical unit *"The chair was comfortable but dirty and its color was awful."* is processed by the following way:
  comfortable (+1) + dirty (-3) + awful (-1): P = -3

- the lexical unit *"It is not good idea."* is processed by the following way:
  good (+1) first dictionary, not (-2) second dictionary: P = 1*[1+(-2)] = -1

- the lexical unit *"The text processing is very decent."* is processed by the following way: decent (+1) first dictionary, very (+1) second dictionary:
  P = 1*[1+1)] = +2.

The question is, how long the length of processed lexical unit should be. We decided to represent the lexical units by n-grams. The value of n was fixed on 4 in an experimental way. Shorter value of n can separate negation or intensifier from related word. Consequently, we have designed new versions of n-gram approach. The first version has created 4-grams applying shift by one position. Another one was based on comparison of polarity value obtained by 4-grams and polarity value obtained by 1-grams.

7 Evaluation and Testing

The implementations of various versions of opinion classification were tested within several experiments.

First, the basic version with static coefficient (called "Static coefficient" in Table 1) was tested on the set of discussion contributions from the portal http://www.mobilmania.sk (a discussion thread related to reviews of the LGKU990 mobile telephone). This set of contributions contained 1558 words within 236 lexical units and the used classification dictionary had 27 positive words, 27 negative words, 10 negations and 11 intensifications. The resulting precision of these tests was 0.78, which was an average calculated from precision of positive contributions classification (higher value 0.86) and precision of negative ones (lower value 0.69).

The next versions were the version using the dynamic coefficient, which was set as an average length of each sentence (called "Dynamic coefficient 1" in the Table 1), the version using the dynamic coefficient determined as half length of each sentence (called "Dynamic coefficient 2" in the Table 1) and the version using the hybrid approach (called "Hybrid" in the Table 1) were tested on 50 reviews (25 positive reviews and 25 negative ones) obtained from the page http://recenzie.sme.sk. The classification dictionary contained 150 words in all categories (positive words, negative words, negations and intensifications). The results of negative contributions classification were surprisingly high as can be seen in Table 1 (0.84, 0.88 and 0.84) – higher than precision of positive contributions classification (0.76, 0.80 and 0.80). It may be caused by the fact, that the number of negative contributions was the same as the number of positive ones. Usually, the precision of negative contributions is lower because the number of available negative contributions is fairly lower.

Lastly, the version based on n-grams was tested in two independent experiments (called "n-grams 1 and n-grams 2" in Table 1). The first experiment was performed on a set of 42 contributions with 2350 words. These contributions were extracted from discussion forums available on http://www.mojandroid.sk (a discussion thread related to reviews of the mobile telephones HTC One X and HCT One S) and http://www.pocitace.sme.sk (a discussion thread related to reviews of two products Asus Transformer Prime TF201 and Asus Transformer Pad TF300T). The second experiment (n-grams 2) was performed on a set of 71 contributions (4341 words) from the portal http://tech.sme.sk (a discussion thread related to reviews of the telephone Samsung Galaxy S4) and from the portal http://www.mojandroid.sk (a discussion thread related to reviews of the telephones HTC ONE and Samsung Galaxy S4). These two tests of n-grams application were provided using classification

dictionary with 44 positive words, 46 negative words, 3 kinds of negations and 10 intensifications. The achieved precision of classification of positive contributions in the first experiment (N-grams 1) was quite satisfactory (0.83) but in the second experiment (N-grams 2) was not very high (0.76). A different situation was in the case of negative contributions classification. The achieved precision in the first experiment was only 0.57 and in the second experiment even worse 0.42, what is very low value. Maybe, an extremely low number of available negative contributions within these two last experiments was the cause of such disappointed results.

Table 1. Precision of testing results of various implementation versions.

Version	Positive	Negative	Average precision
Static coefficient	0.86	0.69	0.78
Dynamic coefficient 1	0.76	0.84	0.80
Dynamic coefficient 2	0.80	0.88	0.84
Hybrid	0.80	0.84	0.82
N-grams 1	0.83	0.57	0.70
N-grams 2	0.76	0.42	0.59

The best average precision was achieved by three implementation versions using the dynamic coefficient. These results are interesting, because these versions had better results for processing negative contributions than for processing positive contributions. This fact is not common within the existing opinion classification applications. Usually, a processing of negative contributions is less precise. The basic version using static coefficient and version based on n-grams are more common from this point of view, because they achieved lower precision within negative contributions while they had quite high precision within positive contributions processing. The higher value of precision than 0.80 (80 percentage) seems to be very good in comparison with the average precision of human advices (about 70 percentage). An important factor is also the language which is used to present the classified opinions. Although languages have some similar features, the complexity of the task is evaluated on the basis of the emergent expression of the given particular language.

Our modified application achieves relatively better results - higher precision (0.84, 0.82 using dynamic coefficient) against [11]. The term "relatively" is used because two kinds of testing results achieved on two different input data sets were compared. They present many tests with resulting performance up to 80 percentages (within more than 100 experiments only two results are higher than 80 percentage). On the other hand, our tests were not so complex as in [11] and therefore our resulting precision could be lower when using their corpuses of reviews. Our results seem to be better (once again relatively) than results (up to 73 percentages) in [12]. This comparison is relative because of compared results were achieved on different input contributions and because no cross validation was used. On the other hand, performance achieved in [2] is higher than performance of our applications.

8   Conclusions

The paper introduced variety of approaches to solving the problem of opinion classification to positive or negative polarity. It also described five various original methods of opinion classifications and it also introduced the test results of implementations of the presented methods.

The novelty of our approach is, that our dictionary was generated directly from a web discussion, which increases the precision of opinion analysis of the given discussion and so the dictionary can easily accommodate some misspelled words as well.

The achieved precision of classification of positive contributions within all presented versions of our application was approximately 80 percentages. It is not so bad in the comparison with other existing applications. On the other hand, precision of classification of negative contributions using n-grams is too low. For the future, we would like to test the introduced version of opinion classification application on larger corpus of discussion contributions.

There is a possibility to upgrade version based on n-grams to achieve higher precision within processing of negative contributions. There is a need to extend the classification dictionary. This version should be enriched by techniques for processing also contributions, which contain only neutral words, but their context is positive or negative. The techniques for processing of irony and ambiguity should be included too. The research in the field of opinion classification has big importance for the future. A successful application of opinion classification can be very helpful in the process of decision making.

References

1. Catena, A., Alexandrov, M., Ponomareva, N.: Opinion Analysis of Publications on Economics with a Limited Vocabulary of Sentiments. International Journal on Social Media - MMM: Monitoring, Measurement, and Mining, Vol. 1, No. 1, 20-31 (2010)
2. Choi, Y., Cardie, C.: Learning with Compositional Semantics as Structural Inference for Subsentential Sentiment Analysis. Proc. of the EMNLP 2008, Conference on Empirical Methods in Natural Language Processing, 793-801 (2008)
3. Ding, X., Liu, B., YuA, P.: Holistic Lexicon-Based Approach to Opinion Mining. Proc. of the Int. Conf. on Web Search and Web Data Mining WSDM'2008, New York, NY, USA, 231-240 (2008)
4. GO, A.: Twitter Sentiment Classification using Distant Supervision. Stanford University, http://cs.stanford.edu/people/alecmgo/papers/ TwitterDistantSupervision09.pdf.
5. Heer, J., Boyd, D.: Vizster: Visualizing Online Social Networks. Proceedings of the IEEE Symposium on Information Visualization INFOVIS'2005, Washington, USA, 5-13 (2005)

6. Kaurova, O., Alexandrov, M., Ponomareva, N.: The Study of Sentiment Word Granularity for Opinion Analysis (a Comparison with Maite Taboada Works). International Journal on Social Media - MMM: Monitoring, Measurement, and Mining, Vol. 1, No. 1, 45-57 (2010)

7. Koncz, P., Paralič, J.: An Approach to Feature Selection for Sentiment Analysis. Proc. of the INES 2011 - 15th International Conference on Intelligent Engineering Systems, Poprad, 2011, 357-362, ISBN 978-142448956-5 (2011)

8. Lukáč, G., Butka, P., Mach, M.: Semantically-enhanced Extension of the Discussion Analysis Algorithm in SAKE. In: SAMI 2008, 6th International Symposium on Applied Machine Intelligence and Informatics : January, 2008, Herľany, Slovakia, 241-246 (2008)

9. Mach, M., Lukáč, G.: A Dedicated Information Collection as an Interface to Newsgroup Discussions, IIS 2007 - 18th international conference on Information and Intelligent Systems, September 12-14, 2007,Varazdin, Croatia, 163-169, ISBN 978-953-6071-30-2 (2007)

10. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. Foundation and Trends in Information Retrieval, Vol.2, No.1-2, 1-135 (2008)

11. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, Vol. 37, No. 2, 267-307 (2011)

12. Thelwall, M., Buckley, K., Paltoglou, G, Cai, D., Kappas, A.: Sentiment Strength Detection in Short Informal Text. Journal of the American Society for Information Science and Technology, Vol. 61, No. 12, 2010, 2544-2558 (2010)

13. Stanford Classifier. Stanford University, http://nlp.stanford.edu/software/classifier.shtml