Using SentiWordNet for multilingual sentiment analysis

CONFERENCE PAPER · MAY 2008

DOI: 10.1109/ICDEW.2008.4498370 · Source: DBLP

1 AUTHOR:

Kerstin Denecke
University of Leipzig

Using SentiWordNet for Multilingual Sentiment Analysis

Kerstin Denecke

Research CenterL3S
Appelstrasse 9a, D-30167 Hannover, Germany
denecke@l3s.de

Abstract – **This paper introduces a methodology for determining polarity of text within a multilingual framework. The method leverages on lexical resources for sentiment analysis available in English (SentiWordNet). First, a document in a different language than English is translated into English using standard translation software. Then, the translated document is classified according to its sentiment into one of the classes "positive" and "negative". For sentiment classification, a document is searched for sentiment bearing words like adjectives. By means of SentiWordNet, scores for positivity and negativity are determined for these words. An interpretation of the scores then leads to the document polarity. The method is tested for German movie reviews selected from Amazon and is compared to a statistical polarity classifier based on n-grams. The results show that working with standard technology and existing sentiment analysis approaches is a viable approach to sentiment analysis within a multilingual framework.**

I. Introduction

Opinion mining concerns with the opinion an author of a document expresses. This analysis of sentiments provokes several challenges: Among other things, it has to be determined whether a document or section thereof is subjective or objective and whether the opinion expressed is positive or negative. Furthermore, the strength of the expressed opinion and the target of an opinion have to be found out. Due to the richness of human language, its large expressiveness and ambiguities the problem of sentiment classification is nontrivial.

The sentiment of a text may be crucial in several applications, like

- mining and summarizing customer, book and movie reviews,
- analyzing political opinions,
- classifying blog posts and comments.

Even though, there are several applications that are also interesting in a multilingual context, most of the research in this area has focused on processing documents in English to date.

In this paper, we investigate methods to automatically determine the polarity of sentences in a multilingual framework. The methods rely on resources available for English. Specifically, through experiments the question will be answered whether it is possible to determine polarity without language-specific lexicon of sentiment-bearing terms and without having language specific training data available. Within a multilingual framework, training material is sparse. Therefore, polarity is determined by means of a lexicon-based approach in the proposed methods where the lexicon is defined for English terms.

The remainder of this paper is structured as follows. After providing an overview on approaches to polarity analysis, above all in a multilingual framework (section II), the approaches used in this paper to determine polarity on sentence and document level are introduced (section III). In section IV, the results of first experiments are presented. The paper will conclude with a discussion of the proposed approaches and with remarks on further steps in developing a multilingual sentiment analysis tool.

II. Related Work

There are mainly two approaches to sentiment classification, i.e. approaches based on lexical resources and natural language processing and approaches employing machine learning algorithms. Some approaches relevant for this paper are described in the following.

A. Multilingual Sentiment Analysis

Sentiment Analysis within a multilingual context offers several challenges. Statistical approaches require training material which is normally sparse for different languages or is even unavailable. On the other hand, lexical approaches necessitate language specific lexical and linguistic resources. Generating these resources is very time consuming and requires often manual work.

According to our knowledge, there are mainly two approaches that are relevant in the context of multilingual sentiment analysis. *Mihalcea et al*. compare in [3] a corpus-based approach and a lexicon-based approach to multilingual subjectivity analysis (subjective vs. objective). Within the lexicon-based approach, a target-language subjectivity classifier is generated by translating an existing lexicon. The corpus-based approach builds a subjectivity-annotated corpus for the target language through projection. A statistical classifier is trained on the resulting corpus.

Ahmad et al. [4] pursue a local grammar approach for sentiment classification within a multilingual framework (English, Arabic, and Chinese). Domain-specific keywords are selected by comparing the distribution of words in a domain-specific document to the distribution of words in a general language corpus. Words less prolific in a general language corpus are considered to be keywords. The context of each keyword (i.e., its surrounding words) helps to produce collocation patterns. By these local grammar patterns sentiment bearing phrases are extracted.

The described approaches do not determine the sentiment of a text: the approach of *Mihalcea et al.* distinguish subjective from objective text. *Ahmad et al.* only detect sentiment bearing phrases. The polarity of the text remains hidden. The approach presented in this paper aims to determine the sentiment of texts within a multilingual framework by using the same (i.e., monolingual) processing resources for different languages.

B. Sentiment Analysis based on SentiWordNet

The approach described in this paper is based on SentiWordNet, a lexical resource for opinion mining. In SentiWordNet (http://sentiwordnet.isti.cnr.it/), to each synset of WordNet, a triple of polarity scores is assigned i.e., a positivity, negativity and objectivity score. The sum of these scores is always 1. For example the triple {0, 1, 0} (positivity, negativity, objectivity) is assigned to the synset of the term “bad”. The sum of all scores of this synset is 1.

SentiWordNet has been created automatically by means of a combination of linguistic and statistic classifiers. It has been applied in different opinion-related tasks, i.e. for subjectivity analysis and sentiment analysis with promising results.

Devitt and Ahmad [5], have already used this resource for sentiment polarity detection in financial news, but in a monolingual context. Their lexicon-based approach relies on SentiWordNet that is applied for quantifying positive and negative sentiments.

The linguistic rule-based approach of *Chaumartin* [6] uses SentiWordNet in combination with WordNet Affect to detect emotion and valence values for words in headlines, but only in a monolingual context.

In [7], subjective adjectives are extracted from SentiWordNet. These adjectives are then used as features for estimating the probability that a document contains opinion bearing expressions.

However, SentiWordNet has not been applied within a multilingual context by now. This will be focused in this paper.

III. METHODS

For classifying sentences of documents regarding their sentiments, a processing pipeline comprising three steps is established (Fig. 1). First, the language of a document is determined using language models. Then, if the document language is not English, the document is translated into English by means of standard translation software. Finally, the translated document is classified into a sentiment class (positive or negative). The processing steps are described in more detail in the following sections.

A. *Language classification*

The language of a document is determined by means of the LingPipe Language Identification Classifier. LingPipe (http://www.alias-i.com/lingpipe) provides a set of open source java libraries for several natural language processing tasks.

The LingPipe Language Identifier treats language identification as classification problem. For each language that has to be identified, a set of training texts is used. For training material, LingPipe refers to the Leipzig Corpora Collection (http://corpora.uni-leipzig.de/) that consists of text corpora in 15 different languages.

The classifier for language identification learns the distribution of characters per language using language models. A language model assigns a probability to a sequence of words $P(w_{1..n})$ by means of a probability distribution. It aims to predict the probability of natural word sequences: Word sequences that actually occur achieve a high probability whereas those sequences that never occur get a low probability.

An n-gram model approximates these probabilities by assuming that the only words relevant to predicting $P(w_i|w_1 \ldots w_{i-1})$ are the previous n-1 words:

$$P(w_i|w_1 \ldots w_{i-1}) = P(w_i|w_{i-n+1} \ldots w_{i-1})$$

Like text classifiers that attempt to identify attributes which distinguish documents in different categories, language models also attempt to capture such regularities.

For text classification using an n-gram language model, a language model for each category is trained based on a set of training data. To classify a new (unknown) document, a language model is calculated and compared to the trained language models. The category of the language model which is most similar to that of the unknown document is assigned.

For language classification with LingPipe, n-grams of size eight are used. Character-based language models are successfully used for different tasks in text classification, e.g. spam detection [10]. More details on language models are given by *Carpenter B.* in [14] and *Peng F* [15].

B. Language standardisation

For mapping blogs and reviews of different languages to English, standard translation software is used. Each document of a different language than English is translated automatically into English. The translated document is considered to be a “correct” translation; it is not corrected by a human.

For our experiments, we are using **PROMT** e**X**cellent **T**ranslation (**XT**) Technology to translate documents into English. PROMT (http://www.e-promt.com) is one of the world leading provider of natural language processing technologies and provides translations for German, English Spanish, French, Portuguese, Italian and Russian, even though for perfectly translated documents some manual rework of the translation would be necessary [9]. According to several quality assessments, PROMT translation software ensures the highest possible quality and efficiency of translation processes to date.

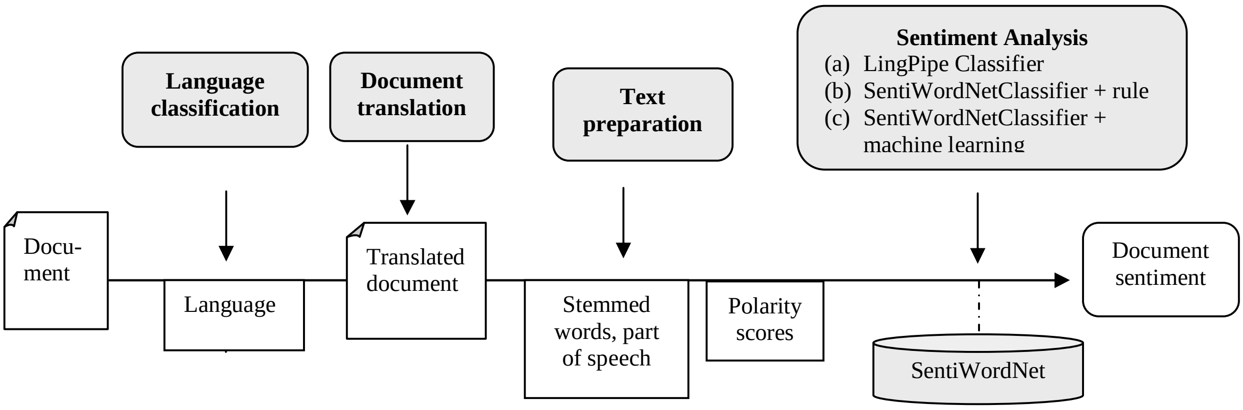

Fig. 1: Processing pipeline for sentiment analysis

C. Sentiment classification

For classifying documents according to their sentiment (positive, negative), three different approaches were implemented and evaluated:

a) LingPipe Classifier,
b) SentiWordNet Classifier with classification rule,
c) SentiWordNet Classifier with machine learning.

(a) LingPipe Polarity analysis

LingPipe's text classification algorithms (see III.A) can also be used for polarity detection. A character-based language model is learned by means of training material for "positive" and "negative" documents. Unknown documents are classified by means of the (learned) language model for "positive" and "negative" documents.

While training material in different languages is sparse, the LingPipe Classifier is trained on documents in English even if the text that has to be classified is written in a different language than English. Based on the language model trained for each class ("positive", "negative"), documents that have been translated into the target language (English) are classified.

(b, c) Polarity Analysis with SentiWordNet

The other two approaches to sentiment analysis introduced in this paper are based on SentiWordNet (see above). The document polarity is calculated from the sentence polarities. Therefore, a document is firstly dissected into sentences. Then, for each sentence the polarity is determined by means of SentiWordNet. Each word is stemmed and tagged, stop words are removed. If a stemmed word belongs to one of the word classes "adjective", "verb" or "noun", it is looked up in SentiWordNet. The scores of corresponding synsets are collected.

Provided that *n* corresponding synsets for a word *A* with part of speech *o* exist, a sentiment score triple ($score_{pos}$, $score_{neg,}$ $score_{obj}$) is determined by means of:

$$(1) \quad score_{pos}(A) = \frac{1}{n}\sum_{i=1}^{n} score_{pos}(i)$$

$$(2) \quad score_{neg}(A) = \frac{1}{n}\sum_{i=1}^{n} score_{neg}(i)$$

$$(3) \quad score_{obj}(A) = \frac{1}{n}\sum_{i=1}^{n} score_{obj}(i)$$

with $score_{\{pos|neg|obj\}}(i)$ = {positivity | negativity |objectivity}- score of synset i for term *A* and *n* = number of synsets of term *A*. The scores of all terms within a sentence are in turn added and divided by the number of terms with scores (example 1).

Example 1:
Considering the sentence "The film, however, is all good." that is stemmed, stop words are removed. The resulting string is "film good". These two words are searched in SentiWordNet for corresponding synsets.
For the input term "film", SentiWordNet provides only synsets with positivity = 0, negativity = 0, objectivity = 1.
For the input term "good", SentiWordNet contains 33 synset entries. After summing up the different scores and dividing the sum by the number of synsets, the resulting values are: positivity = 0.57, negativity = 0.03, objectivity = 0.4.
The sentence score triple results from summing up the score triple of each term and dividing each score by the number of considered terms. The resulting triple for this sentence is: positivity = 0.285, negativity = 0.015, objectivity = 0.7.

The mentioned procedure results in a triple of {positivity, negativity, objectivity}-values for each sentence. To determine a score-triple for a document, the sentence score triples are added and normalized by the number of sentences (in the same manner as in formulas 1-3).

The scores are used in two different ways to classify a document "positive" or "negative". One approach (b) applies a classification rule according to which each document whose positivity score is larger than or equal to the negativity score is classified as “positive”. Otherwise it is considered “negative”.

A different approach (c) is based on machine learning algorithms: The score triple for each document provides the input for a Simple Logistic classifier that has been implemented in the WEKA package [11]. The classifier uses three features: the positivity, negativity and objectivity scores that have been calculated for each document in advance. Based on this, the Simple Logistic Classifier determines the polarity of a document. In the given context, this classifier performs better than the other classifier provided by WEKA.

Even if an objectivity score is calculated during the processing, it is not further considered currently.

IV. Evaluation

To evaluate the proposed approaches, especially regarding multilingual sentiment analysis, several experiments have been performed. Experiment 1 (section IV.B) aims to compare the quality of the introduced methods to sentiment analysis for English documents. Then in section IV.C, the performance of the methods for documents in a different language (German) is determined.

For evaluation purposes, the quality is measured in precision and recall. Recall is a measure of “completeness”, precision is a measure of “cleanness”.

A. Evaluation material

For training and testing the classifiers, different corpora are used. The multi-perspective question answering (MPQA) corpus consists of 535 news articles from 187 different foreign and U.S. news sources and has been used as experimental data by many previous researchers of opinion analysis [1]. The documents were manually annotated at expression level with a variety of subjectivity information. The corpus is available at http://www.cs.pitt.edu/mpqa/databaserelease/.

For training purposes, movie reviews drawn from the IMDB's archive of rec.arts.movies.reviews are used. LingPipe refers on its webpages to these reviews for sentiment classification. 1000 positive and 1000 negative reviews in English are available.

For testing the algorithms with documents in a different language than English, movie reviews from Amazon.de have been collected. The corpus consists of 100 positive and 100 negative movie reviews in German. A movie review from this corpus has been selected as “positive” if four or five stars were assigned by the reviewer. It has been classified as “negative” if the movie got one or two stars. Reviews with three stars were removed from the corpus for this evaluation.

From the linguistic point of view, these three corpora are very different not only with regard to their domain. Whereas the movie reviews in German and English are sometimes very short, the MPQA news articles are quite long. The reviews are written in common speech or even in slang. They may contain writing errors. The authors also express opinions by using a particular style of writing (e.g. if the author writes “Schaaaade” (compared to the correct spelling “Schade” (“What a pity!”) he reinforces his opinion by a different spelling). Compared to this, the news articles are written in correct (formal) English.

B. Experiments on MPQA corpus

The first set of experiments on opinion analysis was carried out on the MPQA corpus. While the MPQA corpus is annotated at expression level, the sum of all the annotated expressions within a sentence was calculated and used to define an overall sentiment at sentence level. If the majority of the annotated expressions within a sentence was “positive” (“negative”) the sentence was considered “positive” (“negative”). In case of a balanced number of “positive” and “negative” expressions within a sentence, it was considered “neutral”. “Neutral” sentences remained unconsidered within this evaluation.

This first evaluation aimed to determine the performance of the different classification methods for English documents. The experiments run at sentence level. 250 positive and 250 negative sentences were selected randomly from the MPQA corpus as evaluation material. The sentences were classified by means of the three described approaches.

The LingPipe Classifier (a) and the machine learning based SentiWordNet Classifier (c) were trained on 1000 positive and 1000 negative full text movie reviews from the IMDB's movie review archive. Precision and recall were calculated for positive and negative sentences (see Table 1).

TABLE 1: EVALUATION RESULTS FOR SENTIMENT CLASSIFICATION ON THE MPQA CORPUS

	positive		**negative**	
	Precision	**Recall**	**Precision**	**Recall**
Classifier (a)	52%	67%	54%	40%
Classifier (b)	55%	70%	75%	32%
Classifier (c)	61%	68%	64%	56%

Positive documents are classified with better recall than negative documents throughout all three algorithms. On the other hand, the precision in classifying negative documents is slightly better than those for positive documents. Algorithm (b) provides a very low recall value for classification of negative documents. The best overall accuracy of 62% is achieved by algorithm (c). The results suggest that positive and negative sentences can be classified best with the algorithm based on SentiWordNet that uses machine learning techniques for classification (c).

Even though the training data and the test data belong to different domains (movie reviews versus news articles), the classification results correspond to results of other published approaches. Better results may probably be achieved by using a training set and a test set that come from the same domain. A detailed discussion of errors is given in section 5.

C. Experiments on movie reviews in German

The performance of the algorithms for a different language is analyzed by means of the Amazon.de corpus (see above).

The classifier (a) and (c) were trained on the same (IMDB) training corpus as in IV.B because a training corpus in German was not available.

The documents were processed according to the described procedures (a), (b) and (c) (see section III). Recall and precision values were calculated (see Tab. 2).

TABLE 2: CLASSIFICATION RESULTS FOR GERMAN MOVIE REVIEWS

	positive		negative	
	Precision	**Recall**	**Precision**	**Recall**
Classifier (a)	80%	28%	55%	92%
Classifier (b)	55%	98%	89%	16%
Classifier (c)	65%	74%	68%	57%

The results differ significantly between the different classifiers. Whereas the LingPipe Classifier (a) provides high precision and low recall values for "positive" documents, the rule-based SentiWordNet Classifier performs for the "negative" documents in that manner. Classifier (a) performs on positive documents as classifier (b) on negative documents and vice versa. The machine-learning-based SentiWordNet Classifier (c) achieves in this evaluation the best accuracy of 66% (Table 3). The recall values of algorithm (c) are for positive and negative documents significantly higher than those of algorithm (a) and (b). The results and sources of error are discussed in the next chapter.

TABLE 3: ACCURACY OF THE DIFFERENT SENTIMENT CLASSIFIER FOR ENGLISH AND GERMAN DOCUMENTS

	English	**German**
Classifier (a)	54%	59%
Classifier (b)	51%	58%
Classifier (c)	62%	66%

V. DISCUSSION

The proposed methods classify news documents in English according to their polarity with an accuracy between 51% and 62%. Movie reviews in German have been slightly better classified with accuracies between 58% and 66%. The results suggest that the accuracy of the different methods does not depend on the processed language and domain. Classifier (a) and (b) perform with a similar accuracy of around 52% for English and 58% for German documents. Above all, the recall values differ significantly for evaluation 1 and 2. These results suggest that translation of documents influences the recall values: While classifying German documents the recall decreased. But in the same time, the precision values increases.

The low recall values of the LingPipe Classifier may due to the training material: The language models for LingPipe were built up using a training corpus in English as a training corpus in German was not available. Other studies [12] have shown that sentiment analysis is a domain-specific task. Therefore, better results could probably be achieved by taking test and training material from the same domain.

There were five situations in particular that were responsible for a significant portion of the errors in the SentiWordNet Classifiers:

- Negated structures are not considered by the SentiWordNet-Classifiers .
- Translation errors or missing translations of words as well as
- writing errors (e.g. "Schaaade" ("What a pity!")) ensure missing sentiment scores for terms.
- Ambiguities and different meanings of a synset in SentiWordNet respectively are not resolved.
- The meaning of words is not considered in context.

The low accuracy of algorithm (b, the SentiWordNet based classifier with fix classification rule) regarding classifying negative documents, is traced to its limited ability to recognize negative texts in a corpora.

Ambiguities are currently not considered by the approach: All synsets that match a specific term are used to calculate a sentiment score triple. To avoid this, the SentiWordNet synsets of an ambiguous term can be reduced to relevant SentiWordNet synsets based on the surrounding terms that can be classified uniquely.

However, the quality for sentiment classification (positive, negative) by means of the approach to multilingual sentiment analysis introduced in this paper is comparable to results of monolingual approaches. The approach analyzed in [12] achieves a maximum accuracy of 68% for sentiment classification on movie reviews. This accuracy is achieved by algorithm (c) in this paper.

Kim and Hovy introduce in [16] a lexicon-based approach to sentiment analysis comparable to the introduced algorithms. They manually collect sentiment-bearing words from WordNet and expand this list by collecting synonyms from WordNet. Within an evaluation, the authors apply their method to 50 emails in German that are translated into English in advance. The evaluation results of the approaches in this paper, in particular the results of classifier (c), are comparable to those in the evaluation of *Kim and Hovy*. Whereas their method achieves a better recall on negative documents, our approach performs better on positive documents (Table 4). This might be due to the composition of the list of sentiment-bearing words or to the test material. The accuracy of our approach is slightly higher than that of the *Kim and Hovy* approach.

TABLE 4: RESULTS OF THE APPROACH OF KIM AND HOVY ON GERMAN EMAILS COMPARED TO CLASSIFIER (C)

	positive		negative	
	Precision	**Recall**	**Precision**	**Recall**
Classifier (c)	65%	74%	68%	57%
Kim & Hovy	72%	40%	55%	80%

The difference between the approach of *Kim and Hovy* and the methods introduced in this paper is the lexical resource the systems are based on: *Kim and Hovy* generate a list of sentiment bearing words with large manual effort. The approach here uses SentiWordNet as lexical resource. Furthermore, the classification methods differ: In this paper, a rule-based classifier and a statistical classifier are compared. *Kim and Hovy* use a classification rule for polarity classification.

In [1], an approach to sentiment analysis is introduced and evaluated on the MPQA corpus that has been used in our experiments, too. That approach is based on a set of ten features (e.g. word token, word polarity, whether it is negated). It achieves precision and recall values for positivitiy classification of 67% and 63% and for negativity classification 73% and 82%. According to this, their method performs better for detecting negative sentences than the approaches compared in this paper. A main reason for this is that negations remain unconsidered by the approaches here. Recognizing negated structures and considering them in sentiment classification would be a potential extension of our approach.

An additional feature for algorithm (c) could be a polarity score calculated as proposed in by *Kale et al.* in [13] that considered the frequencies of positive and negative expressions within a sentence. In our approaches, the frequencies of positive and negative terms remain unconsidered at the moment.

The main benefit of the approach presented in this paper is the use of SentiWordNet as lexical resource. The processing of documents in different languages is based on this resource and it existed already.

VI. CONCLUSIONS

In this paper, an approach to multilingual sentiment classification was described. It has been shown that combining a suite of existing technologies, standard translation software and existing sentiment analysis resources for English achieves good performance in classifying texts according to their sentiment (positive, negative). The accuracy in polarity classification achieves 66%. SentiWordNet has proven a reliable resource for sentiment analysis in a multilingual context.

In future evaluations the methods will be analyzed within a larger test set and for several languages. A comparison of the evaluation results of the introduced methods with results of a sentiment analysis method designed for only one language will allow additional statements regarding the generality of the proposed approach. The SentiWordNet Classifier could be enriched by WordNetAffect as additional resource.

Furthermore, the linguistic / lexicon-based approach can be enhanced with other features, like those introduced in [1]. The question whether syntactical features can help in this approach still has to be considered.

Apart from classifying sentences into “positive” or “negative”, “subjective” and “objective” sentences can be distinguished. Furthermore, more comprehensive statements regarding the sentiment of sentences and documents are possible by means of the calculated polarity triples (e.g., whether a sentence is strongly or weakly positive).

In future research, the topic detection and the detection of the opinion holder within a multilingual context will be analyzed. In this way, the polarity of a document regarding a specific topic can be determined (instead of calculating one global document polarity). In addition, the problem of multilingual opinion analysis will be considered in other domains, e.g. in analyzing entries in medical blogs.

REFERENCES

[1] Wilson T, Wiebe J, Hoffmann P: *Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis*. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing. Vancouver, Canada, 2005, 347-354

[2] Esuli A, Sebastiani F: *SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining*. In: Proceedings of LREC 2006 – 5th Conference on Language Resources and Evaluation

[3] Mihalcea R, Banal C, Wiebe J: *Learning Multilingual Subjectivity Language via Cross-Lingual Projections.* Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, 2007

[4] Ahmad K, Cheng D, Almas Y: *Multi-lingual Sentiment Analysis of Financial News Streams*. In: Proc. of the Second Workshop on Computational Approaches to Arabic Script-based Languages, Linguistic Society of America, Linguistic Institute, Stanford University, 2007, 1-12

[5] Devitt A, Ahmad K: *Sentiment Polarity Identification in Financial News: A Cohesion-based Approach*. Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, June 2007, 984-991

[6] Chaumartin FR: UPAR7: *A knowledge-based system for headline sentiment tagging.* Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval), Prague, June 2007, 422-435

[7] Zhang E, Zhang Y: UCSC *on TREC 2006 Blog Opinion Mining.* TREC 2006 Blog Track, Opinion Retrieval Task

[8] Niu Y, Zhu XD, Li JH, Hirst G: *Analysis of polarity information in medical text*. In: Proceeings of the AMIA Symposium, Washington, D.C., 2005, 570-574

[9] Promt: *A Brief Guide to PROMT Machine Translation Technology.* http://www.e-promt.com/en/technology/pdf/E_Guide_PROMT_MT_Technology.pdf/, last access: 10/10/2007

[10] Fuchun P, Dale S, Shaojun W: *Language and task independent text categorization with simple language models.* In: Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, Edmonton, 2003, 189-196

[11] Witten I, Frank E: *Data Mining: Practical Machine Learning Tools and Techniques.* Morgan Kaufmann Publishers, Amsterdam. 2nd edition, 2005

[12] Turney PD: *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews.* Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Philadelphia, 2002

[13] Kale A *et al.: Modeling Trust and Influence in the Blogosphere Using Link Polarity.* In: Proceedings of the International Conference on Weblogs and Social Media (ICWSM 2007), March 2007

[14] Carpenter B: *Scaling high-order character language models to gigabytes.* In: Proceedings of the 2005 Association for Computational Linguistics Software Workshop, 2005

[15] Peng F, Schuurmans D: *Combining Naïve Bayes and n-Gram Language Models for Text Classification.* ECIR, 2003

[16] Kim SM, Hovy E*: Identifying and Analyzing Judgement Opinions.* Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, New York, 2006, 200-207