

# Comparable English-Russian Book Review Corpora for Sentiment Analysis

Taras Zagibalov<sup>1</sup> and Katerina Belyatskaya<sup>2</sup> and John Carroll<sup>3</sup>

**Abstract.** We present newly-produced comparable corpora of book reviews in English and Russian. The corpora are comparable in terms of domain, style and size. We are using them for cross-lingual experiments in document-level sentiment classification. Quantitative analyses of the corpora and the language differences they exhibit highlight a number of issues that must be considered when developing systems for automatic sentiment classification. We also present experiments with a sentiment classification system applied to the corpora. The results suggest that differences in the way that sentiment is expressed in the two languages lead to large difference in sentiment classification accuracy.

## 1 INTRODUCTION

Automatic classification of document sentiment (and more generally extraction of opinion from text) has recently attracted a lot of interest. One of the main reasons for this is the importance of such information to companies, other organizations, and individuals. Applications include marketing research tools that help a company analyse market or media reaction towards their brands, products or services, or search engines that help potential purchasers make an informed choice about a product they are considering buying.

Most extant sentiment classification systems use approaches based on supervised machine learning, which require substantial manually-produced or -curated resources including texts annotated at the document level and below, sentiment term dictionaries and thesauri, and some level of language analysis.

There are a number of publicly available sentiment-annotated corpora, such as MPQA [15], and Pang and Lee’s Movie Review corpus [8]. However, most of these corpora consist just of English text. As for other languages, we are aware of only one publicly available corpus, of Chinese product reviews [20]. There are other corpora designed for cross-lingual evaluations, but these seem not to be publicly available (for example the NTCIR MOAT corpora of English, Japanese and Chinese [12]).

As part of an on-going research effort in sentiment analysis, we have designed and built comparable corpora of book reviews in English and Russian, which we are making publicly available, in the expectation that they will contribute to

research in cross-lingual sentiment processing<sup>4</sup>. The Russian corpus is probably the first sentiment-annotated resource in that language.

In this paper, as well as describing the corpora and quantifying various relevant aspects of them, we analyse some important language-specific and domain-specific issues that would be likely to impact on automatic sentiment processing. We also describe experiments with standard machine learning sentiment classification technique applied to the corpora.

The paper is structured as follows. Section 2 surveys related work in sentiment classification. Section 3 describes the corpora. Section 4 presents experiments with the corpora and Section 5 concludes.

## 2 RELATED WORK

Most work on sentiment classification has used approaches based on supervised machine learning. For example, Pang et al. [9] collected movie reviews that had been annotated with respect to sentiment by their authors, and used this data to train supervised classifiers. A number of studies have investigated the impact on classification accuracy of different factors, including choice of feature set, machine learning algorithm, and pre-selection of the segments of text to be classified. For example, Dave et al. [2] experiment with the use of linguistic, statistical and n-gram features and measures for feature selection and weighting. Pang and Lee [8] use a graph-based technique to identify and analyse only subjective parts of texts. Yu and Hatzivassiloglou [18] use semantically oriented words for identification of polarity at the sentence level. Most of this work assumes binary classification (positive and negative), sometimes with the addition of a neutral class (in terms of polarity, representing lack of sentiment).

Turney [13] carried out an early and influential study into unsupervised sentiment classification. The approach starts from two ‘seed’ words and builds lists of positive and sentiment vocabulary from large amounts of text using a technique based on pointwise mutual information. For sentiment classification of movie reviews the approach achieves a relatively modest 65% accuracy (although reviews of automobiles are classified with 84% accuracy). Turney attributes this discrepancy in accuracy between domains to the much more complex structure of movie reviews. Popescu and Etzioni [10] extend the approach, applying hand-made rules, linguistic information and WordNet resources. Kobayashi et al. [4] employ

<sup>1</sup> University of Sussex, email: T.Zagibalov@sussex.ac.uk

<sup>2</sup> Siberian Federal University, email: e.o.belyatskaya@gmail.com

<sup>3</sup> University of Sussex, email: J.A.Carroll@sussex.ac.uk

<sup>4</sup> The corpora are available for download from <http://www.informatics.sussex.ac.uk/users/tz21/>.

an iterative semi-automatic approach to extracting opinion-bearing expressions, although this requires human input at each iteration. Unsupervised and semi-supervised techniques may offer the promise of overcoming domain dependence since they do not require training data in order to be applied to a new domain. Wiebe and Riloff [14] present an unsupervised sentence-level subjectivity classifier that uses an extensive set (about 8000) of rules (subjectivity clues). Li, Zhang and Sindhvani [5] used labelled documents to adjust a hand-built sentiment lexicon to a domain. The extensive use of knowledge (rule or lexicons) make these approaches language-dependent.

An alternative approach to overcoming domain dependence is presented by Aue and Gamon [3], who attempt to solve the problem of the absence of large amounts of labelled data by customizing sentiment classifiers to new domains using training data from other domains. Blitzer et al. [1] investigate domain adaptation for sentiment classifiers using structural correspondence learning.

There has been little previous work on applying sentiment analysis to languages with scarce relevant language resources. A notable exception is the work towards producing cross-lingual subjectivity analysis resources from English data by Mihalcea et al. [7]. They use a parallel corpus to adjust a subjectivity lexicon translated from English to Romanian. Other multilingual opinion mining work (in English, Japanese and Chinese) was carried out by Zagibalov and Carroll ([19] and [21]), using techniques requiring limited manual input to classify newswire documents with respect to subjectivity and to extract opinion holders and targets.

A number of studies include development of linguistic resources for sentiment analysis. The text corpora are quite often annotated by a several annotators to produce different kinds of annotation. For example, Read [11] developed an annotation scheme with about 30 different tags that closely follows the Appraisal Theory [6]. Wilson and Wiebe [16] developed a detailed annotating scheme for expressions of opinions, beliefs, emotions, sentiment and speculation. To ensure annotation robustness, the authors calculate inter-annotator agreement. Another approach uses tags produced by authors ('self tagged') of the documents included to the corpus [2].

### 3 THE CORPORA

The English and Russian book review corpora consist of reader reviews of science fiction and fantasy books by popular authors. The reviews were written in 2007, ensuring that the language used is current.

The Russian corpus consists of reviews of Russian translations of books by popular science-fiction and fantasy authors, such as S. King, S. Lem, J.K. Rowling, T. Pratchett, R. Salvatore, J.R.R. Tolkien as well as by Russian authors of the genre such as S. Lukianenko, M. Semenova and others. The reviews were published on the website [www.fenzin.org](http://www.fenzin.org).

The English corpus comprises reviews of books by the same authors if available. If some of the authors were not reviewed on the site or did not have enough reviews, they were substituted with other writers of the same genre. As a result the English corpus contains reviews of books such as: S. Erickson (*Guardians of the Moon, Memories of Ice*), S. King (*Christine, Duma Key, Gerald's Game, Different Season* and others), S. Lem (*Solaris, Star Diaris of Iyon Tichy, The Cy-*

*briad*), A. Rise (*Interview with the Vampire, The Tale of the Body Thief* and others), J.K. Rowling (*Harry Potter*), J.R.R. Tolkien (*The Hobbit, The Lord of the Rings, The Silmarillion*), S. Lukyanenko (*The Night Watch, The Day Watch, The Twilight Watch, The Last Watch*), and some others. The reviews were published on the website [www.amazon.co.uk](http://www.amazon.co.uk).

We annotated each review as 'POS' if positive sentiment prevails or 'NEG' if the review is mostly negative based on the tags assigned by reviewers, but moderated where the tag was obviously incorrect. Each corpus consists of 1500 reviews, half of which are positive and half negative. The annotation is simple and encodes only the overall sentiment of a review, for example:

[TEXT = POS]

Hope you love this book as much as I did. I thought it was wonderful!

[/TEXT]

English reviews contain a mean of 58 words (the mean length for positive and negative reviews being almost the same). Positive Russian reviews have a mean length of only 30 words; negative reviews are slightly longer, at 38 words (see Table 1). It is not possible to compare these figures directly across the languages as they have different grammar structures which makes English more 'wordy' as it has function words (articles, auxiliary verbs) which are almost completely absent in Russian.

Russian, being a synthetic language, has a lot of forms of the same lemma. This results in a large number of distinct word forms: the corpus contains a total of 13472 word forms, with 6589 (42%) in positive reviews and 8993 (58%) in negative. The total number of words in the corpus is 50745, which means that every word form was used a little more than 3 times on average. The English corpus has only 7489 word forms in the whole corpus, 4561 (47%) in positive reviews, and 5098 (53%) in negative. The re-use of word forms in English is much higher: every word form was used 9 times on average (the total number of word in the corpus is 87539). These figures suggest that the Russian reviewers used a richer vocabulary for expressing *negative* opinion (compared to the number of unique words used in Russian positive reviews) than English reviewers.

Further evidence of the different ways in which people distinguish sentiment polarity in Russian compared with English is the distribution of lengths of positive and negative reviews. The Russian corpus has a large number of short reviews (less than 50 words) with a median of 15 words for positive reviews and 10 words for negative reviews. Apart from the language-specific differences mentioned above that partly account for the smaller number of words in Russian documents, there is a clear difference from English reviews in terms of length. The English reviews feature a more or less equal number of documents of different lengths (mostly in the range 15 to 75). The prevalence of short reviews in the Russian corpus, together with the rich morphological variation, may lead to data sparseness which would be a problem for many current sentiment classification techniques.

Although both of the sites from which the reviews were collected feature review-ranking systems (e.g. one to ten stars), many reviewers did not use the system or did not use it properly. For this reason all of the reviews were read through

	Mean tokens POS	Mean tokens NEG	Total types POS	Total types NEG
English	58	58	7349	8014
Russian	30	38	9290	12309

**Table 1.** Overall quantitative measures of the English and Russian corpora.

and hand-annotated. There were a lot of re-occurring short reviews like: Хорошо (*Good*); Интересная книга (*Interesting book*); Супер! (*Superb!*); Нудятина!! (*Boring!!*); Ниже среднего (*Below average*); Awesome!; Amazing!; The best book I've ever read!; Boring, and so on. These reviews were added to the corpus only once. Also both sites had a number of documents which did not have any direct relation to book reviewing, such as advertisements, announcements and off-topic postings. Such texts were excluded as irrelevant.

The documents that were included in the corpora were not edited or altered in any other way.

### 3.1 Ways of Expressing Sentiments

To better understand the difference between the English and the Russian corpora, we have investigated the means used to express opinion and how this may impact on automatic sentiment classification<sup>5</sup>.

Sentiment can be expressed at different levels in a language, from lexical and phonetic levels up to the discourse level. This range is reflected in the corpora (see Tables 2 and 3). As the Tables show, the two languages express sentiment in slightly different ways. English makes heavy use of adjectives to express sentiment (this class of words is used to express sentiment in a third of all documents). In contrast, Russian uses verbs as often as adjectives to express sentiment (both of these classes are used in about quarter of all reviews) and makes more use of nouns (expressing sentiment in 15% of all documents compared to 11% in English). The Russian corpus also demonstrates a tendency to combine different ways of expressing sentiments in a document: the total number of uses of different ways in the English corpus is 4083 compared to 4716 in Russian, which means that given equal number of reviews for each language, Russian reviews tend to have more different ways of expressing sentiment per document.

	Syntactic	Lexical				Phonetic
		Verb	Adj	Noun	Other	
Positive	432	312	708	225	325	12
Negative	367	389	652	238	407	16
Total	799	701	1360	463	732	28

**Table 2.** Ways of expressing sentiment in the English Book Review Corpus (numbers of documents).

#### Lexical Level

<sup>5</sup> All the numerical data presented below comes from manual counting and is not represented in the corpus annotation.

	Syntactic	Lexical				Phonetic
		Verb	Adj	Noun	Other	
Positive	417	492	648	374	367	27
Negative	475	578	567	334	394	43
Total	892	1070	1215	708	761	70

**Table 3.** Ways of expressing sentiment in the Russian Book Review Corpus (numbers of documents).

**Adjectives** Adjectives are the most frequent means of expressing opinions in both languages, closely followed by verbs in the Russian corpus. 1215 Russian reviews use adjectives to express sentiment and 1070 reviews use verbs. In the English corpus there are 1360 reviews that use adjectives, but only 701 use verbs to express opinion.

Apart from adjectives, which are recognised as the main tool for expressing evaluation, other parts of speech are also often used in this function, most notably verbs and nouns. The English reviews also feature adverbials and both languages also use interjections.

**Verbs** As observed by some researchers, opinions delivered by verbs are more expressive compared to opinions expressed in other ways. This is explained by the fact that a verb's denotation is a situation and the semantic structure of the verb reflects linguistically relevant elements of the situation described by the verb. Appraisal verbs not only name an action, but also express a subject's attitude to an event or fact. Consider the following examples:

- (1) I truly loved this book, and I KNOW you will, too!
- (2) понравилось, научная фантастика в хорошем исполнении  
I liked it, it's science fiction in a very good implementation

The English verbs *loved* and *liked* describe a whole situation which is completed by the time of reporting it. This means that a subsequent shift in sentiment polarity is all but impossible:

- (3) \*I truly loved this book, but it turned out to be boring.

**Nouns** Nouns can both identify an object and provide some evaluation of it. But nouns are less frequently used for expressing opinion compared to verbs. Nonetheless in the Russian corpus, nouns were used more than in the English corpus. There are 708 Russian reviews that have opinions expressed by nouns, however, only 463 English reviews made use of a noun to describe opinion. The most frequent such nouns used in Russian reviews are чудо (*miracle*), классика (*classics*), шедевр (*masterpiece*), гений (*genius*), прелесть (*delight*), бред (*nonsense*), мура (*raspberry*), жвачка (*mind-numbing stuff*), ерунда (*bugger*).

**Phonetic Level** Although the corpora consist of written text and do not have any speech-related mark-up, some of the review authors used speech-related methods to express sentiment, for example:

- (4) A BIG FAT ZEEROOOOOOOOOOOOOOO for M.A
- (5) Ну что сказать... чепуха... ЧЕ-ПУ-ХА.  
What should I say... boloney... ВО-ЛО-НЕУ

Another way to express opinion in Russian is based on the use of a sub-culture language, Padonky. This sociolect has distinctive phonetic and lexical features that are distant from 'standard' Russian (both official and colloquial). For example, a phrase usually used to express negative attitude to an author about his book:

- (6) Аффтор, выпей ЙАДУ  
(lit) Autor, drink some POIZON

Padonky is close to some variants of slang (corresponding in English to expressions such as *u woz*, *c u soon* etc.), however it is more consistent and is used quite often on the Web.

**Sentence Level** Sentence-level means of expressing sentiment (mostly exclamatory clauses, imperatives or rhetorical questions) is slightly more frequent in the Russian corpus than in the English: 892 and 799 respectively. The distribution of positive and negative sentiments realised at the sentence level is opposite in the two corpora: syntactic means are used more frequently in negative reviews in Russian but they are more frequent in positive reviews in English.

One particularly common sentiment-relevant sentence-level phenomenon is the rhetorical question. This is a question only in form, since it usually expresses a statement. For example:

- (7) И откуда столько восторженных отзывов? Коробит от крутости главных героев  
Why are there so many appreciative reviews? The 'coolness' of the main characters makes me sick
- (8) Что же такого пил/принимал/нюхал автор, чтобы написать такое?  
What did the author drink / eat / sniff to write stuff like that?

Some 'borderline' cases like the following are also used to express sentiment:

- (9) Интересно, кто-нибудь дотянул хотя бы до середины? Лично я - нет.  
I wonder if anyone managed to get to the middle? I failed.

Considering imperatives, the review author is telling their audience 'what to do', which is often to read a book or to avoid doing so.

- (10) Run away! Run away!

- (11) Pick up any Pratchett novel with Rincewind and re-read it rather than buying this one

- (12) Читать однозначно.  
Definitely should read.

- (13) Читать !!!!!!!!!!! ВСЕМ  
Read!!!!!!!!!! EVERYONE

Another way of expressing sentiment by means of syntactic structure is exclamatory clauses, which are by their very nature affective. This type of sentence is widely represented in both corpora.

- (14) It certainly leaves you hungering for more!
- (15) Buy at your peril. Mines in the bin!

**Discourse Level** Some of the means of sentiment expression are quite complex and difficult to analyse automatically:

- (16) И это автор вычислителя и леммингов? ...  
so this author calculator and lemmings? ...  
НЕ ВЕРЮ! Садись, Громов, два.  
(DO)NOT BELIVE! sit gromov two  
So is this the author of The Calculator and of The Lemmings? ... Can't believe it! Sit down, Gromov, mark 'D'!

This short review of a new book by Gromov, the author of the popular novels *The Calculator* and *The Lemmings*, consists of a rhetorical question, an exclamatory phrase and an imperative. All of these means of expression are difficult to process. Even the explicit appraisal expressed by utilising a secondary school grade system is problematic as it requires specialised real-word knowledge. Otherwise the numeral 'two'<sup>6</sup> has nothing to do with appraisal per se.

The example below also features an imperative sentence used to express negative sentiment. This review also lacks any explicit sentiment markers. The negative appraisal is expressed by the verbs 'stab' and 'burn' that only in this context show negative attitude.

- (17) Stab the book and burn it!

**Discussion** The reviews in English and in Russian often use different means of expressing sentiment, many of which are difficult (if at all possible) to process automatically. Often opinions are described through adjectives (86% of reviews contain adjectives). The second most frequent way of expressing sentiment is through verbs (59% of reviews have sentiment-bearing verbs). Less frequent is the noun, in 39% of reviews. Sentence-level and discourse-level sentiment phenomena are found in 56% of reviews. 3% of reviews contain phonetic phenomena.

<sup>6</sup> Russian schools use a 5-grade marking system, with 5 as the highest mark. Thus 2 can be thought of as equivalent to 'D'.

### 3.2 Issues that may Affect Automatic Processing

One of the features of web content not mentioned above is a high level of **mistakes and typos**. Sometimes authors do not observe the standard rules on purpose (for example using sociolects, as outlined above). For example, in the corpora 52% of all documents contain spelling mistakes in words that have sentiment-related meaning. The English corpus is less affected as authors do not often change spelling on purpose and use contractions that have already become conventional (e.g. *wanna*, *gonna*, and *u*). However the number of spelling mistakes is still high: 48% of reviews contain mistakes in sentiment-bearing words. The number of misspelled words in the Russian corpus is higher, at 58%.

Of course, a spelling error is not always fatal for automatic sentiment classification of a document, since reviews usually have more sentiment indicators than just one word. However, as many as 8% of the reviews in both corpora have all of their sentiment bearing words misspelled. This would pose severe difficulties for automatic sentiment classification.

Another obstacle that makes sentiment analysis difficult is **topic shift**, in which the majority of a review describes a different object and compares it to the item under review. The negative review below is an example of this:

- (18) Дочитала с трудом. Ничего интересного с точки зрения информации. Образец интеллектуального детектива – романы У.Эко. И читать приятно, и глубина философии, и в историческом плане познавательно. А в эстетическом отношении вообще выше всяких похвал.  
Hardly managed to read to the end. Nothing interesting from the point of view of information. An example of intellectual detective stories are novels by U.Eco. It's a pleasure to read them, and (they have) deep philosophy, and quite informative from the point of view of history. And as for aesthetics it's just beyond praise.

The novel being reviewed is not the one being described, and all the praise goes to novels by another author. None of the positive vocabulary has anything to do with the overall sentiment of the review's author towards the book under review.

Other reviews that are difficult to classify are those that describe some positive or negative aspects of a reviewed item, but in the end give an overall **sentiment of the opposite direction**. Consider the following positive review:

- (19) Сюжет довольно обычен, язык изложения прост до безобразия. Много грязи, много крови и смерти. Слишком реально для сказки коей является фэнтези. Но иногда такие книги читать полезно, ибо они описывают неприглядную реальность.  
The plot is quite usual, the language is wickedly simple. A lot of filth, a lot of blood and death. Too true-to-life for a fairy-tale, which a fantasy genre actually is. But it is useful to read such books from time to time, as they depict ugly reality.

The large number of negative lexical units may mislead an automatic classifier to a conclusion that the review is negative. The three issues described above are present in approxi-

mately one third of all reviews in the corpora. This suggests that a sentiment classifier using words as features could only correctly classify around 55–60% of all reviews.

This performance may be even worse for the Russian corpus as many its reviews feature very unexpected ways of expressing opinion. Unlike most of the English reviews, in which a reviewer simply gives a positive or negative appraisal of a book backing it with some reasoning and probably providing some description and analysis of the plot, Russian reviews often contain **irony, jokes, and use non-standard words and phrases**, making use of a variety of language tools, as illustrated in the following examples:

- (20) Скушнаа. дошёл до бегства ГГ в мир Януса, и внезапно понял (я), что гори он (ГГ) хоть синим пламенем  
Booorin'. got to the (episode of) GG fleeing to the world of Janus, and suddenly (I) realised that (lit.) let it (GG) burn with blue flames ( $\approx$  I do not at all care about GG)
- (21) Я эту муть не покупал. Shift+del.  
I didn't buy this garbage. Shift+del.

Since there are more reviews of this kind in the Russian corpus than in the English, it is very likely that a Russian sentiment classifier would have lower accuracy.

## 4 EXPERIMENTS

We used Naïve Bayes multinomial (NBm) and a Support Vector Machine classifiers<sup>7</sup> to investigate performance of standard supervised classifiers on the two corpora. The feature sets were the lexical units extracted from the relevant corpora. We extracted all words from the corpora but did not process them in any way (no stemming or lemmatisation). 15582 words were extracted from the Russian corpus and 9659 words were found in the English book reviews. The evaluation technique is 10-fold cross-validation.

Corpus	NBm			SVM		
	P	R	F	P	R	F
English book reviews	0.88	0.88	0.88	0.84	0.84	0.84
Russian book reviews	0.81	0.81	0.81	0.78	0.78	0.78

**Table 4.** Supervised classification results (Precision, Recall and  $F_1$ , 10-fold cross-validation)

Table 4 show the results of supervised classification, Russian review classification being 6-7 percentage points worse the results obtained from the English corpus.

## 5 CONCLUSION

In this paper we presented comparable corpora of English and Russian book reviews, providing the research community with a resource that can be used for cross-lingual sentiment classification experiments. We examined language-specific features

<sup>7</sup> We used WEKA (http://www.cs.waikato.ac.nz/ml/weka) [17]

of the reviews that are relevant to sentiment classification and showed that sentiment in different languages is expressed in slightly different ways, covering all levels of the language: from phonetic to discourse. The experiments suggest that these differences have an impact on the accuracy of a standard, supervised sentiment classification technique.

In future work, we intend to investigate in more depth which specific characteristics of different languages lead to differences in sentiment classification accuracy, using sentiment-annotated corpora of English, Russian, Chinese and Japanese.

## References

- [1] John Blitzer, Mark Dredze, and Fernando Pereira, ‘Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification’, in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics.*, pp. 440–447, Prague, Czech Republic, (June 2007). Association for Computational Linguistics.
- [2] Kushal Dave, Steve Lawrence, and David M. Pennock, ‘Mining the peanut gallery: Opinion extraction and semantic classification of product reviews’, in *Proceedings of the 12th international conference on Information and Knowledge Management*, pp. 519 – 528, Budapest, Hungary, (2003). ACM Press.
- [3] Michael Gamon and Anthony Aue, ‘Automatic identification of sentiment vocabulary: exploiting low association with known sentiment terms’, in *Proceedings of the ACL Workshop on Feature Engineering for Machine Learning in Natural Language Processing*, pp. 57–64. Association for Computational Linguistics, (2005).
- [4] Nozomi Kobayashi, Kentaro Inui, Yuji Matsumoto, Kenji Tateishi, and Toshika Fukushima, ‘Collecting evaluative expressions for opinion extraction’, *Natural Language Processing–IJCNLP 2004*, **13**(12), 596–605, (December 2004).
- [5] Tao Li, Yi Zhang, and Vikas Sindhwani, ‘A Non-negative Matrix Tri-factorization Approach to Sentiment Classification with Lexical Prior Knowledge’, in *Proceeding of Association for Computational Linguistics*, number August, pp. 244—252, Morristown, NJ, USA, (2009). Association for Computational Linguistics.
- [6] J.R. Martin and Peter Robert Rupert White, *The language of evaluation: Appraisal in English*, Palgrave Macmillan, 2005.
- [7] Rada Mihalcea, Carmen Banea, and Janyce M Wiebe, ‘Learning multilingual subjective language via cross-lingual projections’, in *976 Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, volume 45, pp. 976—983, Prague, Czech Republic, (2007).
- [8] Bo Pang and Lillian Lee, ‘A sentimental education: Sentiment analysis using subjectivity summarization based on Minimum Cuts’, in *the 42nd Annual Meeting on Association of Computational Linguistics*, Barcelona, Spain, (2004).
- [9] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan, ‘Thumbs up?: sentiment classification using machine learning techniques’, in *Conference on Empirical Methods in Natural Language Processing*, pp. 79—86, (2002).
- [10] Ana-Maria Popescu and Oren Etzioni, ‘Extracting product features and opinions from reviews’, in *Natural Language Processing and Text Mining*, pp. 9–28, Vancouver, Canada, (October 2005). Springer.
- [11] Jonathon Read, David Hope, and John Carroll, ‘Annotating expressions of appraisal in English’, *ACL 2007*, 93, (2007).
- [12] Yohei Seki, David K. Evans, Lun-Wei Ku, Le Sun, Hsin-Hsi Chen, and Noriko Kando, ‘Overview of multilingual opinion analysis task at NTCIR-7’, *Proceedings NTCIR-7, NII, Tokyo*, 185–203, (2008).
- [13] Peter D. Turney, ‘Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews’, in *Annual Meeting of Association of Computational Linguistics*, pp. 417–424, Philadelphia, Pennsylvania, (2002).
- [14] Janyce M Wiebe and Ellen Riloff, ‘Creating subjective and objective sentence classifiers from unannotated texts’, *Computational Linguistics and Intelligent Text Processing*, 486–497, (2005).
- [15] Janyce M Wiebe, Theresa Ann Wilson, and Claire Cardie, ‘Annotating expressions of opinions and emotions in language’, *Language Resources and Evaluation*, **39**(2), 165–210, (2005).
- [16] Theresa Ann Wilson and Janyce M Wiebe, ‘Annotating opinions in the world press’, in *4th SIGdial Workshop on Discourse and Dialogue (SIGdial-03)*, pp. 13–22, (2003).
- [17] I.H. Witten and Eibe Frank, *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann Pub, San Francisco, 2nd edn., 2005.
- [18] Hong Yu and Vasileios Hatzivassiloglou, ‘Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences’, in *Proceedings of EMNLP*, volume 3, pp. 129–136. Association for Computational Linguistics, (2003).
- [19] Taras Zagibalov and John Carroll, ‘Almost Unsupervised Cross Language Opinion Analysis at NTCIR 7’, in *NTCIR-7*, Tokyo, (2008).
- [20] Taras Zagibalov and John Carroll, ‘Automatic Seed Word Selection for Unsupervised Sentiment Classification of Chinese Text’, in *Proceedings of the 22nd International Conference on Computational Linguistics*, pp. 1073—1080, Manchester, United Kingdom, (2008).
- [21] Taras Zagibalov and John Carroll, ‘Multilingual Opinion Holder and Target Extraction using Knowledge-Poor Techniques’, in *Language and Technology Conference*, Poznań, Poland, (2009).