

# MayAnd at SemEval-2016 Task 5: Syntactic and word2vec-based Approach to Aspect-based Polarity Detection in Russian

Vladimir Mayorov and Ivan Andrianov

Institute for System Programming of the Russian Academy of Sciences  
25 Alexander Solzhenitsyn Street  
Moscow, Russia

vmayorov@ispras.ru, ivan.andrianov@ispras.ru

## Abstract

This paper describes aspect-based polarity detection system for Russian, used in aspect-based sentiment analysis task (ABSA) of SemEval-2016 (Task 5, subtask 1, slot 3). The system consists of two independent classifiers: for opinion target expressions and for implicit opinion target mentions. We introduce a set of standard unigram features together with more sophisticated ones: based on sentence syntactic structure and based on lemmas vector representation.

Being applied to Russian restaurant reviews, our system achieved best quality among four participants.

## 1 Introduction

Aspect-based sentiment analysis task (ABSA) drew much attention with the growth of Internet retail sites and facilities for customers to make a free text product review. ABSA systems allow businesses to track reputation of their products even more granularly than before: per aspect. At the same time, customers can be provided with per-aspect sentiment summaries of reviews about concurrent products to simplify their choice.

SemEval-2016 Task 5 (ABSA) organizers (Pontiki et al., 2016) proposed two subtasks: Sentence-level ABSA and Text-level ABSA. First subtask was split to three slots: Aspect Category Detection, Opinion Target Expression, Sentiment Polarity. Training data for the task were available for various languages (English, Russian, etc.) and domains (restaurants, laptops, etc.).

Slot 3 can be described as follows: given a set of reviews labeled with OTEs (opinion target expressions: start-end-entity-attribute quadruples) determine a polarity label for each OTE: positive, negative, neutral (mildly positive or mildly negative sentiment). Additionally polarity labels should be computed for entity-attribute pairs assigned to reviews sentences.

This paper describes our team participation in slot 3 evaluation for Russian language Restaurant domain. The rest of the article is structured as follows: section 2 mentions some recent approaches to aspect-based polarity detection task, section 3 contains our system description, section 4 provides our system evaluation results.

## 2 Related work

According to (Andrianov et al., 2015), aspect-based polarity detection methods can be divided to three categories: classifier-based (Zhang and Lan, 2015; Hamdan et al., 2015; Blinov and Kotelnikov, 2015), neural network-based (Tarasov, 2015) and unsupervised (García Pablos et al., 2015). Most promising approaches according to recent evaluations results (Pontiki et al., 2015; Loukachevitch et al., 2015) are classifier-based methods, especially semi-supervised ones which bring “external knowledge” (word2vec (Mikolov et al., 2013), Brown clusters (Brown et al., 1992), polarity lexicons (Baccianella et al., 2010; Wilson et al., 2005)) into the table. Semi-supervised techniques become extremely useful due to lack of labeled data for the aspect-based sentiment analysis task in general and polarity detection subtask in particular.

### 3 System Description

Our aspect-based polarity detection system consists of two independent parts: OTE polarity classifier and implicit entity mention polarity classifier. Both parts are based on linear SVM with L2 regularization for three classes: positive, negative, neutral (available training set contains the only conflict polarity label, thus it is ignored).

At first, we preprocess review using Texterra (Tur-dakov et al., 2014) system. We perform tokenization, part of speech tagging, lemmatization and syntactic parsing for each review sentence.

All used features can be grouped to several classes:

- opinion target features;
- sentence-level lexical features;
- syntactic-level features.

Some most valuable features are based on word vector representation obtained by word2vec. For word2vec training we used 40140 unlabeled reviews from restoclub<sup>1</sup> resource. All word2vec models were trained with same dimensionality (100) and same window size (10). We tried word2vec with different input data preprocessing approaches: without preprocessing, with lowercased words and with lemmatized words. The last one demonstrates the best result on 8-fold cross-validation (see section 4 for details).

The rest of this section describes each group of features in details.

#### 3.1 Opinion Target Features

Initially we tried to train separate SVM classifiers for each entity-aspect pair. Unfortunately training data contains just a few examples of certain entity-attribute pairs (e.g., *DRINKS#PRICES* are presented in training data only 17 times). Moreover, we suspect that different entities can share most part of sentiment vocabulary for some attributes (e.g., quality).

So, we decided to use a single SVM classifier for all entity-aspect labels and to extract two additional features: opinion target entity label and opinion target attribute label. These features are used in both system parts.

<sup>1</sup>www.restoclub.ru

#### 3.2 Sentence-level Features

As implicit entity mention has no boundaries, we use the whole sentence to extract its features. We use a number of common features, usually exploited in NLP tasks: sentence word unigrams, sentence lemma unigrams, part of speech tag of each word in sentence. Moreover we use frequency of punctuation token unigrams in sentence as a feature.

Described features group is used for implicit entity mention polarity classifier as well as for OTE polarity classifier, because of slight quality improvement.

For OTE polarity classification information about entity mention is also used: we utilize vector representations of lemmatized sentence words. We extract lemmas in window of size 5 (5 words to the left and 5 words to the right from OTE), map them to vector representations and concatenate these representations.

#### 3.3 Syntactic-level features

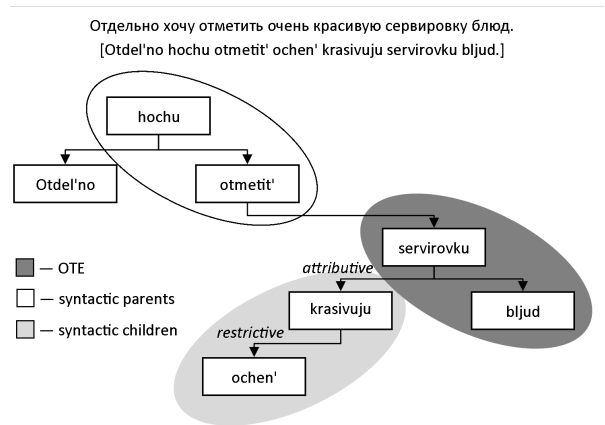


Figure 1: Syntactic dependency parsing for review sentence “I would like to mention very nice table appointments.”

OTE polarity classifier exploits a number of syntactic-level features. Our basic assumption is that OTE polarity is closely related to some dependent words (especially adjectives), which may express sentiment. Our system uses Texterra syntactic dependency parser to produce syntactic tree of input sentence.

Each OTE should match subtree in syntactic tree, which is treated as a single node. In case OTE is not a subtree, syntactic-level features for classifi-

Data source	Preprocessing	Accuracy	Positive F1	Negative F1	Neutral F1
Wikipedia	–	0.8095	0.8907	0.5615	0.2766
	lowercase	0.8089	0.8909	0.5515	0.2820
	lemmatization	0.8113	0.8928	0.5566	0.2930
Restoclub	–	0.8137	0.8937	0.5700	0.2906
	lowercase	0.8158	0.8955	0.5662	0.3086
	lemmatization	<b>0.8189</b>	<b>0.8966</b>	<b>0.5817</b>	<b>0.3271</b>

Table 1: 8-fold cross-validation with different word2vec training settings

ation are abandoned. For instance, Figure 1 depicts syntactic tree extracted from review sentence “I would like to mention very nice table appointments.”. Opinion target expression “Table appointments” is a subtree, so it is treated as a single node.

We start from using OTE parent words as features: we extract a chain of parent words for given OTE (up to 3 words), and use their lemmas vector representations as features. In example syntactic tree (Figure 1) there are two parent words: “otmetit” at depth 1 and “hochu” at depth 2.

Additionally we use syntactic children chains to extract some relevant features. For each word dependent on OTE (in parent-child chains of length up to 2) we use syntactic relations path conjoint with word form, lemma and lemma vector representation as features. In example (Figure 1) syntactic children words are “krasivuju” with relations path “attributive” and “ochen” with relations path “attributive-restrictive”.

## 4 Evaluation

### 4.1 word2vec Training Settings

During system development we tried several word2vec models trained with different settings. We varied data sources (Russian Wikipedia and restaurant reviews obtained from restoclub.ru) and data preprocessing (without preprocessing, lowercased words and lemmatized words). We experimented with 8-fold cross-validation on training data. The best result was obtained with restoclub.ru data source and lemmatization preprocessing (see Table 1).

### 4.2 Features Impact

In this section we try to evaluate each feature group impact. We test quality of our system with all de-

scribed features on test data (using gold labels provided by organizers after evaluation) and using 8-fold cross-validation on train data (as we did before evaluation). Then we consecutively test our system without some group of features. Test results are presented in Table 2.

Features	cv on train data	on test data
All	<b>0.8189</b>	0.7808
All - opinion target	0.8128 (-0.0061)	<b>0.7923 (+0.0123)</b>
All - sentence-level	0.7793 (-0.0396)	0.7038 (-0.0770)
All - syntactic	0.8142 (-0.0047)	0.7777 (-0.0031)
All - word2vec	0.8029 (-0.0160)	0.7661 (-0.0147)

Table 2: Features impact evaluation. Accuracy score of the system with various features sets

In spite of the fact that word lemmas vector representation features are present in sentence-level and syntactic features groups, we additionally test these features impact separately.

As it could be seen, most significant feature groups are sentence level features and lemma vector representation features. However, large impact of sentence-level features group is explained with the fact that implicit entity mention polarity classifier uses a few additional features, and when we don’t use sentence-level features it uses only opinion target features. If we test OTE polarity classifier separately, we obtain 0.8115 accuracy score without sentence-level features versus 0.8287 with all features on cross-validation (0.0172 impact) and 0.7746 versus 0.7860 on test data (0.0114 impact).

### 4.3 Evaluation Results

The official results are presented in Table 3. Our system (MayAnd) official result differs from result presented in Table 2 due to regrettable inadvertence with training dataset. Our system was trained on ini-

tially published data instead of revised dataset that had been published later. Table 2 reflects actual quality of our system (trained on the latest published dataset).

System name	Result (Accuracy)
MayAnd	<b>0.7792</b>
INSIG.	0.7508
IIT-T.	0.7362
Danii.	0.7331
baseline	0.7100

**Table 3:** Evaluation official results

## 5 Conclusion

Aspect-based sentiment analysis task and polarity detection subtask are still at their early days. Current methods are not able to significantly overcome trivial baselines. The following directions of future work can be recommended: collecting more reviews for various languages and domains to improve quality of unsupervised techniques like word2vec; development of sentiment lexicons for non-English languages and various domains.

## Acknowledgments

This work has been supported by the RFBR grant No. 15-37-20375.

## References

- Ivan Andrianov, Vladimir Mayorov, and Denis Turdakov. 2015. Modern approaches to aspect-based sentiment analysis. In *Proceedings of the Institute for System Programming*, volume 27, pages 5–22.
- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204.
- Pavel Blinov and Evgeny Kotelnikov. 2015. Semantic similarity for aspect-based sentiment analysis. In *Proceedings of the 21st International Conference on Computational Linguistics Dialog-2015*, volume 2, pages 36–45.
- Peter F Brown, Peter V deSouza, Robert L Mercer, Vincent J Della Pietra, and Jenifer C Lai. 1992. Class-based n-gram models of natural language. *Computational linguistics*, 18(4):467–479.
- Aitor García Pablos, Montse Cuadros, and German Rigau. 2015. V3: Unsupervised aspect based sentiment analysis for semeval2015 task 12. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 714–718, Denver, Colorado, June. Association for Computational Linguistics.
- Hussam Hamdan, Patrice Bellot, and Frederic Bechet. 2015. Lsislif: Crf and logistic regression for opinion target extraction and sentiment polarity analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 753–758, Denver, Colorado, June. Association for Computational Linguistics.
- Natalia Loukachevitch, Pavel Blinov, Evgeny Kotelnikov, Yuliya Rubtsova, Vladimir Ivanov, and Elena Tutubalina. 2015. Sentirueval: testing object-oriented sentiment analysis systems in russian. In *Proceedings of the 21st International Conference on Computational Linguistics Dialog-2015*, volume 2, pages 12–24.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. Semeval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 486–495, Denver, Colorado, June. Association for Computational Linguistics.
- Maria Pontiki, Dimitrios Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeny Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit. 2016. SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval ’16*, San Diego, California, June. Association for Computational Linguistics.
- Dmitriy Tarasov. 2015. Deep recurrent neural networks for multiple language aspect-based sentiment analysis of user reviews. In *Proceedings of the 21st International Conference on Computational Linguistics Dialog-2015*, volume 2, pages 77–88.
- Denis Turdakov, Nikita Astrakhantsev, Yaroslav Nedumov, Andrey Sysoev, Ivan Andrianov, Vladimir Mayorov, Denis Fedorenko, Anton Korshunov, and Sergey Kuznetsov. 2014. Texterra: A framework for text analysis. *Programming and Computer Software*, 40(5):288–295.
- Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level

sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics.

Zhihua Zhang and Man Lan. 2015. Ecnu: Extracting effective features from multiple sequential sentences for target-dependent sentiment analysis in reviews. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 736–741, Denver, Colorado, June. Association for Computational Linguistics.