

Neural Networks for Sentiment Analysis in Czech

Ladislav Lenc^{1,2} and Tomáš Hercig^{1,2}

¹ Department of Computer Science and Engineering, Faculty of Applied Sciences,
University of West Bohemia, Univerzitní 8, 306 14 Plzeň, Czech Republic

² NTIS—New Technologies for the Information Society, Faculty of Applied Sciences,
University of West Bohemia, Technická 8, 306 14 Plzeň, Czech Republic

nlp.kiv.zcu.cz

{llenc,tigi}@kiv.zcu.cz

Abstract: This paper presents the first attempt at using neural networks for sentiment analysis in Czech. The neural networks have shown very good results on sentiment analysis in English, thus we adapt them to the Czech environment. We first perform experiments on two English corpora to allow comparability with the existing state-of-the-art methods for sentiment analysis in English. Then we explore the effectiveness of using neural networks on four Czech corpora. We show that the networks achieve promising results however there is still much room for improvement especially on the Czech corpora.

Keywords: Sentiment Analysis, Neural Networks, Czech

1 Introduction

The current approaches to sentiment analysis in English explore various neural network architectures (e.g. [1, 2, 3]). We try to replicate the results shown in [1] and adapt the proposed architecture to the sentiment analysis task in Czech – a highly inflectional Slavic language. To the best of our knowledge, neural networks have not been used for the sentiment analysis task in Czech.

The goal of aspect-based sentiment analysis (ABSA) is to identify the aspects of a given target entity and estimate the sentiment polarity for each mentioned aspect, while the general goal of sentiment analysis is to detect the polarity of a text. In this work we will focus on polarity detection on various levels (texts, sentences, and aspects).

In recent years the aspect-based sentiment analysis has undergone rapid development mainly because of competitive tasks such as SemEval 2014 - 2016 [4, 5, 6].

Aspect-based sentiment analysis firstly identifies the aspects of target entity and then assigns a polarity to each aspect. There are several ways to define aspects and polarities.

The definition of the ABSA task from SemEval 2014 distinguishes two types of aspect-based sentiment: aspect terms and aspect categories. The whole task is divided into four subtasks.

- **Aspect Term Extraction (TE)** – identify aspect terms.

Our server checked on us maybe twice during the entire meal.
→ {server, meal}

- **Aspect Term Polarity (TP)** – determine the polarity of each aspect term.

Our server checked on us maybe twice during the entire meal.
→ {server: negative, meal: neutral}

- **Aspect Category Extraction (CE)** – identify (predefined) aspect categories.

Our server checked on us maybe twice during the entire meal.
→ {service}

- **Aspect Category Polarity (CP)** – determine the polarity of each (pre-identified) aspect category.

Our server checked on us maybe twice during the entire meal.
→ {service: negative}

The later SemEval’s ABSA tasks (2015 and 2016) further distinguish between more detailed aspect categories and associate aspect terms (targets) with aspect categories.

The current ABSA task - SemEval 2016 [6] has three subtasks: Sentence-level (SB1), Text-level (SB2) and Out-of-domain ABSA (SB3). The subtasks are further divided into three slots. The following example is from the training data (including the typographical error).

- **1) Aspect Category Detection** – identify (predefined) aspect category – entity and attribute (E#A) pair.

The pizza is yummy and I like the atmospere.
→ {FOOD#QUALITY, AMBIENCE#GENERAL}

- **2) Opinion Target Expression (OTE)** – extract the OTE referring to the reviewed entity (aspect category).

The pizza is yummy and I like the atmospere.
→ {pizza, atmospere}

- **3) Sentiment Polarity** – assign polarity (positive, negative, and neutral) to each identified E#A, OTE tuple.

The pizza is yummy and I like the atmospere.
→ {FOOD#QUALITY - pizza: positive,
AMBIENCE#GENERAL - atmospere: positive}

In this work we will focus on the sentiment polarity task on aspect-level and document-level¹ for Czech and English. In terms of the SemEval 2014 task it is the *Aspect Term Polarity* and *Aspect Category Polarity* (TP and CP) subtasks. In terms of the SemEval 2016 task it is the *Sentence-level Sentiment Polarity* subtask.

Our main goal is to measure the difference between the previous results and the new results achieved by neural network architectures.

2 Related Work

2.1 Sentiment Analysis in Czech

Initial research on Czech sentiment analysis has been done in [7, 8, 9, 10]. However they used only small news datasets and because of the small data size no strong conclusions can be drawn.

The first extensive evaluation of Czech sentiment analysis was done by Habernal et al. in [11]. Three different classifiers, namely Naive Bayes, Support Vector Machines and Maximum Entropy classifiers were tested on large-scale labeled corpora (Facebook posts, movie reviews, and product reviews). In [12] they further experimented with feature selection methods.

Habernal and Brychcin [13] used semantic spaces (see [14]) created from unlabeled data as an additional source of information to improve results. Brychcin and Habernal [15] explored the benefits of the global target context and outperformed the previous unsupervised approach.

The first attempt at aspect-based sentiment analysis in Czech was presented in [16]. This work provides an annotated corpus of 1244 sentences from the restaurant reviews domain and a baseline ABSA model. Hercig et al. [17] extended the dataset from [16], nearly doubling its size and presented results using several unsupervised methods for word meaning representation.

The work in [18] creates a dataset in the domain of IT product reviews. This dataset contains 200 annotated sentences and 2000 short segments, both annotated with sentiment and marked aspect terms (targets) without any categorization and sentiment toward the marked targets. Using 5-fold cross validation on the aspect term extraction task (TE) they achieved 65.79% F-measure on the short segments and 30.27% F-measure on the long segments.

2.2 Neural Networks and Sentiment Analysis

First attempt to estimate sentiment using a neural network was presented in [19]. The authors propose using Active Deep Networks which is a semi-supervised algorithm. The network is based on Restricted Boltzmann Machines. The approach is evaluated on several review datasets containing an earlier version of the movie review dataset created

¹For the English RT dataset and Czech Facebook dataset it can be also called the sentence-level.

by Pang and Lee [20]. It outperforms the state of the art approaches on these datasets.

Ghiassi et al. [21] use Dynamic Artificial Network for sentiment analysis of Tweets. The network uses n -gram features and creates a Twitter-specific lexicon. The approach is compared to Support Vector Machines classifier and achieves better results.

Socher et al. [2] utilizes a Recursive Neural Tensor Network trained on the Stanford Sentiment Treebank (SST). The network is tested on the binary sentiment classification and on the fine-grained (continuous number from 0 to 1) sentiment polarity scale. It outperforms the state of the art methods on both tasks.

A Deep Convolutional Neural Network is utilized for sentiment classification in [3]. Classification accuracies of 48.3% (5 sentiment levels) and 85.7% (binary) on the the SST dataset are achieved.

Several papers propose more general neural networks used for NLP tasks that are tested also on sentiment datasets. One of such methods is presented in [1]. A Convolutional Neural Network (CNN) architecture is proposed and tested on several datasets such as Movie Review (MR) dataset and SST. The tasks were sentiment classification (binary or 5 sentiment levels), subjectivity classification (subjective/objective) and question type classification. It proved state-of-the-art performance on all datasets.

In [22] a Dynamic Convolutional Neural Network is proposed. A concept of dynamic k-max pooling is used in this network. It is tested on sentiment analysis and question classification tasks.

The authors of [23] propose two CNNs for ontology classification, sentiment analysis and single-label document classification. Their networks are composed of 9 layers out of which 6 are convolutional layers and 3 fully-connected layers with different numbers of hidden units and frame sizes. They show that the proposed method significantly outperforms the baseline approaches (bag of words) on English and Chinese corpora.

3 Data

In this work we use two types of corpora:

- Aspect-level for the ABSA task and
- Document-level for the sentiment polarity task.

The properties of these corpora are shown in Table 1. The English Aspect-level datasets come from the SemEval ABSA tasks. Although we show properties of the datasets from previous years, we report results only on the latest datasets from the SemEval 2016.

We do not use the Czech IT product datasets because of its small size and because no results for the sentiment polarity task have been reported using these datasets so far. The Czech Facebook dataset has a label for bipolar sentiment which we discard, similarly to the original publication.

Table 1: Properties of the aspect-level and document-level corpora in terms of the number of *sentences*, *average length* of sentences (number of words), and numbers of *positive*, *negative*, *neutral* and *bipolar* labels.

Aspect-level Sentiment Dataset	Sentences	Average Length	Positive	Negative	Neutral	
English 2016 Laptops train + test	3.3k	14	2.1k	1.4k	0.2k	
English 2016 Restaurants train + test	2.7k	13	2.3k	1k	0.1k	
English 2015 Restaurants train + test	2k	13	1.7k	0.7k	0.1k	
Czech Restaurant reviews	2.15k	14	2.6k	2.5k	1.2k	
Czech IT product reviews short	2k	6	1k	1k	–	
Czech IT product reviews long	0.2k	144	0.1k	0.1k	–	
Document-level Sentiment Dataset	Sentences	Average Length	Positive	Negative	Neutral	Bipolar
English RT Movie reviews	10.7k	21	5.3k	5.3k	–	–
Czech CSFD Movie reviews	91.4k	51	30.9k	29.7k	30.8k	–
Czech MALL Product reviews	145.3k	19	103k	10.4k	31.9k	–
Czech Facebook posts	10k	11	2.6k	2k	5.2k	0.2k

For all experiments we use 10-fold cross validation in cases where there are no designated test and train data splits.

4 System

The proposed sentiment classification system can be divided into two modules. The first one serves for data preprocessing and creates the data representation while the second one performs the classification. The classification module utilizes three different neural network architectures. All networks use the same preprocessing.

4.1 Data Preprocessing and Representation

The importance of data preprocessing has been proven in many NLP tasks. The first step in our preprocessing chain is removing the accents similarly to [12] and converting the text to lower case. This process may lead to loss of some information but we include it due to the fact that the data we use are collected from the Internet and therefore it may contain grammatical errors, misspellings and could be written either with or without accents. Finally, all numbers are replaced with one common token. We also perform stemming utilizing the High Precision Stemmer [24].

The input feature of the neural networks is a sequence of words in the document represented using the one hot encoding. A dictionary is first created from the training set. It contains a specified number of most frequent words. The words are then represented by their indexes in the dictionary. The words that are not in the dictionary are assigned a reserved "Out of dictionary" index. An important issue is the variable length of classified sentences. Therefore, we cut the longer ones and pad the shorter ones to a fixed length. The padding token has also a reserved index. We use dictionary size 20,000 in all experiments. The sentence length was set to 50 in all experiments with document-level sentiment. We set the sequence length to 11 in the aspect-level sentiment experiments.

4.2 CNN 1

This network was proposed by Kim in [1]. It is a modification of the architecture proposed in [25]. The first layer is the embedding one. It learns a word vector of fixed length k for each word. We use $k = 300$ in all experiments. It uses one convolutional layer which is composed of a set of filters of size $n \times k$ which means that it is applied on sequence of n words and the whole word vector (k is the length of the word vector). The application of such filters results in a set of feature maps (results after applying the convolutional filters to the input matrix). Kim proposes to use multiple filter sizes ($n = 3, 4, 5$) and utilizes 100 filters of each size. Rectified linear units (Relu) are used as activation function, drop-out rate is set to 0.5 and the mini-batch size is 50. After this step, a max-over-time pooling is applied on each feature map and thus the most significant features are extracted. The selection of one most important feature from each feature map is supposed to ensure invariance to the sentence length. The max pooling layer is followed by a fully connected softmax layer which outputs the probability distribution over labels. There are four approaches to the training of the embedding layer:

- **1)** Word vectors trained from scratch (randomly initialized)
- **2)** Static word2vec [26] vectors
- **3)** Non-static vectors (initialized by word2vec and then fine tuned)
- **4)** Multichannel (both random initialized and pre-trained by word2vec).

The hyper-parameters of the network was set on the development set from SST-2 dataset². We use identical configuration in our experiments to allow comparability. We implemented only the basic – randomly initialized version of word embeddings. Figure 1 depicts the architecture of the network.

²Stanford Sentiment Treebank with neutral reviews removed and binary labels

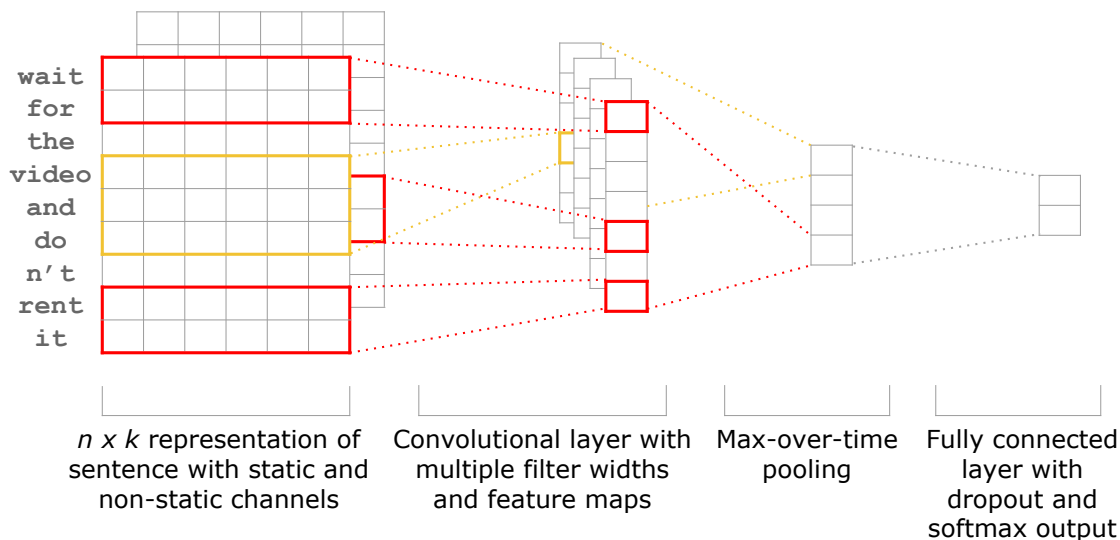


Figure 1: Architecture of the convolutional network CNN1

4.3 CNN 2

The architecture of this network was designed according to [27] where it is successfully used for multi-label document classification.

Contrary to the work of Kim [1] this network uses just one size of the convolutional kernels and not the combination of several sizes. The kernels have only 1 dimension (1D) while Kim have used larger 2 dimensional kernels. It was proven on the document classification task that the simple 1D kernels give better results than the 2D ones.

The input of the network is a vector of word indexes as described in Section 4.1. The first layer is an embedding layer which represents each input word as a vector of a given length. The document is thus represented as a matrix with l rows and k columns where k is the length of embedding vectors. The embedding length is set to 300. The next layer is the convolutional one. We use n_c convolution kernels of the size $l_k \times 1$ which means we do 1D convolution over one position in the embedding vector over l_k input words. The size k is set to 3 (aspect-level sentiment) and 5 (document-level sentiment) in our experiments and we use $n_c = 32$ kernels. The following layer performs max pooling over the length $l - l_k + 1$ resulting in $n_n 1 \times k$ vectors. The output of this layer is then flattened and connected with the output layer containing either 2 or 3 nodes (number of sentiment labels). Figure 2 shows the architecture of the network.

4.4 LSTM

The word sequence is the input to an embedding layer same as for the CNNs. We use the embedding length of 300 in all experiments. The word embeddings are then fed to the recurrent LSTM layer with 128 hidden neurons. Dropout rate of 0.5 is then applied and the final state of the

LSTM layer is connected with the softmax output layer. The network architecture is depicted in Figure 3.

4.5 Tools

We used Keras [28] for implementation of all above mentioned neural networks. It is based on the Theano deep learning library [29]. It has been chosen mainly because of good performance and our previous experience with this tool. All experiments were computed on GPU to achieve reasonable computation times.

5 Experiments

Results on RT movie dataset [30] (10662 sentences, 2 classes) confirm that our implementation works similarly to the original (see Table 2).

Table 2: Accuracy on the English RT movie reviews dataset in %.

Description	Results
Kim [1] randomly initialized	76.1
Kim [1] best result	81.5
CNN1	77.1
CNN2	76.2
LSTM	61.7
Confidence Interval	± 0.8

We further performed evaluation on the current SemEval 2016 ABSA dataset to allow comparison with the current state-of-the-art methods. These results (see Table 3) show that the used neural network architectures are still quite far from the finely tuned state-of-the-art results. However we need to remind the reader that our goal was

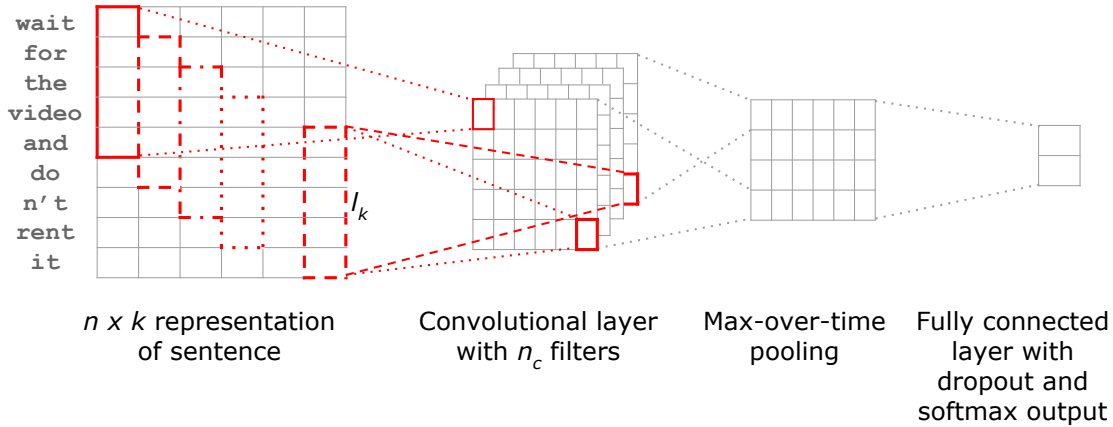


Figure 2: Architecture of the convolutional network CNN2

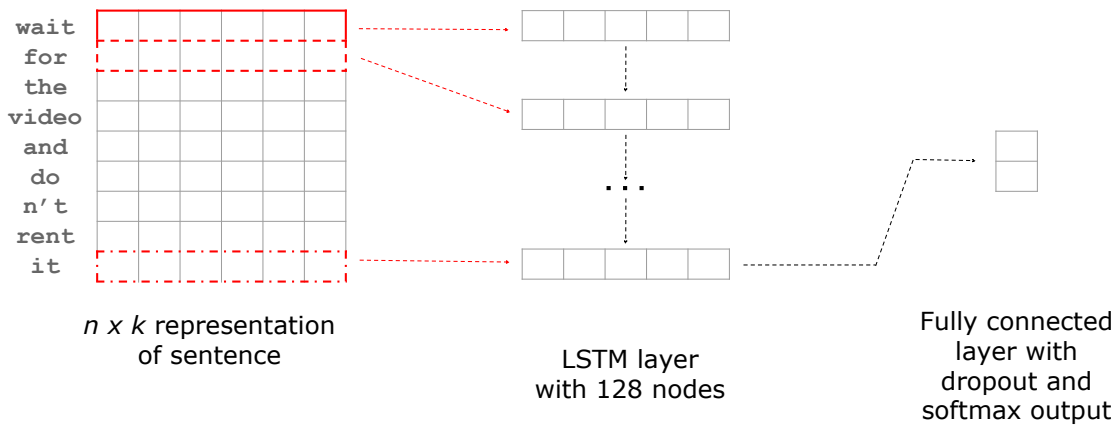


Figure 3: Architecture of the LSTM network

not to achieve the state-of-the-art results, but to replicate network architectures that are used for sentiment analysis in English as well as some networks utilized for other tasks in Czech.

Results on the Czech document-level datasets are shown in Table 4. For the CSFD movie dataset, results are much worse than the previous work. We believe that this is due to the number of words used for representation. We used 50 words in all experiments and it may not suffice to fully understand the review. This is supported by the fact that the global target context [15] helps to improve the results by 1.5%.

We applied three types of neural networks to the term polarity (TP) and class polarity (CP) tasks and evaluated them on the Czech aspect-level restaurant reviews dataset. The results in Table 5 show markedly inferior results compared to the state-of-the-art results 72.5% for the TP and 75.2% for the CP tasks in [17]. Best results are achieved using the combination of words and stems as input.

The inputs of the networks are one-hot vectors created from words in the context window of the given aspect

term. We used five words in each direction of the searched aspect term resulting in window size 11. We do not use any weighting to give more importance to the closest words as in [17].

For statistical significance testing, we report confidence intervals at α 0.05.

CNN1 and CNN2 present similar results although the average best performance is achieved by the CNN2 architecture. The LSTM architecture consistently underperforms, we believe that this is due to the basic architecture model.

6 Conclusion and Future Work

In this work we have presented the first attempts to classify sentiment of Czech sentences using a neural network. We evaluated three architectures.

We first performed experiments on two English corpora mainly to allow comparability with existing work for sentiment analysis in English.

Table 3: Accuracy on the English SemEval 2016 ABSA datasets in %.

Description	Restaurants	Laptops
SemEval 2016 best result	88	82
SemEval 2016 best constrained	88	75
CNN1	78	68
CNN2	78	71
LSTM	72	68
Confidence Interval	± 3	± 3

Table 4: F-measure on the Czech document-level datasets in %.

Description	CSFD Movies	MALL Products	Facebook Posts
Supervised Machine Learning [11]	78.5	75.3	69.4
Semantic Spaces [13]	80	78	-
Global Target Context [15]	81.5	-	-
CNN1 stemmed	70.8	74.4	68.9
CNN2 stemmed	71.0	75.5	69.4
LSTM stemmed	70.2	73.5	67.6
Confidence Interval	± 0.3	± 0.2	± 1.0

Table 5: Accuracy on the Czech aspect-level restaurant reviews dataset in %. *W* denotes words, *S* stems and *W+S* the combination of these inputs.

Description \ Features	Term Polarity			Class Polarity		
	W	S	W+S	W	S	W+S
CNN1	65	66	67	65	66	68
CNN2	64	65	66	67	68	69
LSTM	61	62	62	65	65	64
Confidence Interval	± 2	± 2	± 2	± 2	± 2	± 2

We have further experimented with three Czech corpora for document-level sentiment analysis and one corpus for aspect based sentiment analysis. The experiments proved that the tested networks don't achieve as good results as the state-of-the-art approaches. The most promising results were obtained when using the CNN2 architecture. However, regarding the confidence intervals, we can consider the performance of the architectures rather comparable.

The results show that Czech is much more complicated to handle when determining sentiment polarity. This can be caused by various properties of Czech language that differ from English (e.g. double negative, sentence length, comparative and superlative adjectives, or free word order). Double or multiple negatives are grammatically correct ways to express negation in Czech while in English double negative is not acceptable in formal situations or in writing. Thus the semantic meaning of sentences with double or multiple negatives is hard to determine. In English comparative and superlative forms of adjectives are created by adding suffixes³ while in Czech suffixes and prefixes are used. Informal texts can contain mixed irregular adjectives with prefixes and/or suffixes thus making it

harder to determine the semantic meaning of these texts. The free word order can also cause difficulties to train the models because the same thing may be expressed differently.

However, it must be noted that the compared approaches utilize much richer information than our basic features fed to the neural networks. The neural networks were also not fine-tuned for the task. Therefore we believe that there is much room for further improvement and that neural networks can reach or even outperform the state-of-the-art results.

We consider this paper to be the initial work on sentiment analysis in Czech using neural networks. Therefore, there are numerous possibilities for the future work. The obtained results must be thoroughly analysed to identify cases where the neural networks fail. An interesting experiment would be sentiment analysis on Czech data automatically translated to English. One possible direction of further improvement is utilizing word embeddings to initialize the embedding layer. We also plan to experiment with neural networks on the other two tasks of aspect based sentiment analysis – aspect term extraction and aspect category extraction. Another perspective is to develop new neural network architectures for sentiment analysis.

³excluding irregular and long adjectives

Acknowledgements

This work was supported by the project LO1506 of the Czech Ministry of Education, Youth and Sports and by Grant No. SGS-2016-018 Data and Software Engineering for Advanced Applications. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures".

References

- [1] Kim, Y.: Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882* (2014)
- [2] Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., Potts, C.: Recursive deep models for semantic compositionality over a sentiment treebank. In: *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*. Volume 1631., Citeseer (2013) 1642
- [3] dos Santos, C.N., Gatti, M.: Deep convolutional neural networks for sentiment analysis of short texts. In: *COLING*. (2014) 69–78
- [4] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., Manandhar, S.: SemEval-2014 Task 4: Aspect based sentiment analysis. In: *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, Ireland, Association for Computational Linguistics and Dublin City University (August 2014) 27–35
- [5] Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., Androutsopoulos, I.: Semeval-2015 task 12: Aspect based sentiment analysis. In: *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, Association for Computational Linguistics, Denver, Colorado. (2015) 486–495
- [6] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., Clercq, O.D., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S.M., Eryiğit, G.: SemEval-2016 task 5: Aspect based sentiment analysis. In: *Proceedings of the 10th International Workshop on Semantic Evaluation. SemEval '16*, San Diego, California, Association for Computational Linguistics (June 2016)
- [7] Veselovská, K., Hajič jr., J., Šindlerová, J.: Creating annotated resources for polarity classification in Czech. In Jancsary, J., ed.: *Proceedings of KONVENS 2012, ÖGAI* (September 2012) 296–304 *PATHOS 2012 Workshop*.
- [8] Steinberger, J., Lenkova, P., Kabadjov, M., Steinberger, R., van der Goot, E.: Multilingual entity-centered sentiment analysis evaluated by parallel corpora. In: *Proceedings of the 8th International Conference Recent Advances in Natural Language Processing*. (2011) 770–775
- [9] Steinberger, J., Ebrahim, M., M., E., Hurriyetoglu, A., Kabadjov, M., Lenkova, P., Steinberger, R., Tanev, H., Vázquez, S., Zavarella, V.: Creating sentiment dictionaries via triangulation. *Decision Support Systems* **53** (2012) 689–694
- [10] Veselovská, K., Hajič jr., J.: Why words alone are not enough: Error analysis of lexicon-based polarity classifier for Czech. In: *Proceedings of the 6th International Joint Conference on Natural Language Processing, Nagoya, Japan, Asian Federation of Natural Language Processing, Asian Federation of Natural Language Processing* (2013) 1–5
- [11] Habernal, I., Ptáček, T., Steinberger, J.: Sentiment analysis in Czech social media using supervised machine learning. In: *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Atlanta, GA, USA, Association for Computational Linguistics (June 2013) 65–74
- [12] Habernal, I., Ptáček, T., Steinberger, J.: Supervised sentiment analysis in Czech social media. *Information Processing & Management* **50**(5) (2014) 693–707
- [13] Habernal, I., Brychcín, T.: Semantic spaces for sentiment analysis. In: *Text, Speech and Dialogue*. Volume 8082 of *Lecture Notes in Computer Science*., Berlin, Springer-Verlag (2013) 482–489
- [14] Brychcín, T., Konopík, M.: Semantic spaces for improving language modeling. *Computer Speech & Language* **28**(1) (2014) 192 – 209
- [15] Brychcín, T., Habernal, I.: Unsupervised improving of sentiment analysis using global target context. In: *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, Shoumen, Bulgaria, INCOMA Ltd. (September 2013) 122–128
- [16] Steinberger, J., Brychcín, T., Konkol, M.: Aspect-level sentiment analysis in Czech. In: *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Baltimore, MD, USA, Association for Computational Linguistics (June 2014)
- [17] Hercig, T., Brychcín, T., Svoboda, L., Konkol, M., Steinberger, J.: Unsupervised methods to improve aspect-based sentiment analysis in Czech. *Computación y Sistemas* (in press)
- [18] Tamchyna, A., Fiala, O., Veselovská, K.: Czech aspect-based sentiment analysis: A new dataset and preliminary results. *Proceedings of the 15th conference ITAT 2015* (2015) 95–99
- [19] Zhou, S., Chen, Q., Wang, X.: Active deep networks for semi-supervised sentiment classification. In: *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, Association for Computational Linguistics (2010) 1515–1523
- [20] Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, Association for Computational Linguistics (2002) 79–86
- [21] Ghiassi, M., Skinner, J., Zimbra, D.: Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with applications* **40**(16) (2013) 6266–6282
- [22] Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences. *arXiv*

- preprint arXiv:1404.2188 (2014)
- [23] Zhang, X., LeCun, Y.: Text understanding from scratch. arXiv preprint arXiv:1502.01710 (2015)
 - [24] Brychcín, T., Konopík, M.: Hps: High precision stemmer. *Information Processing & Management* **51**(1) (2015) 68–91
 - [25] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (almost) from scratch. *The Journal of Machine Learning Research* **12** (2011) 2493–2537
 - [26] Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. (2013) arXiv preprint arXiv:1301.3781.
 - [27] Lenc, L., Král, P.: Deep neural networks for Czech multi-label document classification. In: *International Conference on Intelligent Text Processing and Computational Linguistics*, Konya, Turkey (April 3 - 9 2016)
 - [28] Chollet, F.: Keras. <https://github.com/fchollet/keras> (2015)
 - [29] Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J., Warde-Farley, D., Bengio, Y.: Theano: a cpu and gpu math expression compiler. In: *Proceedings of the Python for scientific computing conference (SciPy)*. Volume 4., Austin, TX (2010) 3
 - [30] Pang, B., Lee, L.: Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In: *Proceedings of the ACL*. (2005)