UWB at SemEval 2014 and 2016

Tomáš Brychcín[1], Tomáš Hercig[2], Lukáš Svoboda[2], and Michal Konkol[1]

[1]NTIS – New Technologies for the Information Society,
University of West Bohemia, Univerzitní 8, 306 14 Plzeň
[2]Department of Computer Science and Engineering,
University of West Bohemia, Univerzitní 8, 306 14 Plzeň

{brychcin, tigi, svobikl, konkol}@kiv.zcu.cz

Abstract. International Workshop on Semantic Evaluation (SemEval) is an ongoing series of evaluations of NLP (Natural Language Processing) algorithms, organized by Association for Computational Linguistics (ACL), the international scientific society which hold the major NLP conferences. The evaluations are intended to explore different aspects of meaning in a natural language. The results of NLP algorithms are compared with human judgments. The submitted systems from research teams across the world are compared in terms of performance. Our research team actively participates in the SemEval exercises. This paper summarizes our results in the area of semantic textual similarity and aspect-based sentiment analysis. In 2014 and 2016 our systems were among the best performing in both mentioned tasks.

Key words: SemEval, Semantic Textual Similarity, Aspect-based Sentiment Analysis, Distributional Semantics

1 Introduction

Natural language processing (NLP) is a progressive research field of computer science, artificial intelligence, and computational linguistics. Challenges in NLP involve human-computer interaction and the natural language understanding.

During the last years, NLP research has focused mainly on the semantic analysis, which investigates the ways, how to represent and how to automatically infer the meaning of a text. It has become the core NLP task and can be seen at prestigious conferences as the main topic. A better semantic models result in better performance of the particular NLP tasks (named entity recognition [4], language modeling [5], sentiment analysis [2, 3], document classification, summarization, stance detection, machine translation, and many others).

International Workshop on Semantic Evaluation (SemEval) is a shared task for evaluation of semantic models. It is organized by Association for Computational Linguistics (ACL), the society which holds the major NLP conferences. The results of semantic models are compared with human judgments. Our NLP research team actively participates in the SemEval tasks. This paper describes our results achieved at

J. Steinberger, M. Zíma, D. Fiala, M. Dostal, M. Nykl (eds.)
Data a znalosti 2017, Plzeň, 5. - 6. října 2017, pp. 111-115.

SemEval 2014 and 2016, concretely, in the tasks semantic textual similarity (Section 3) and aspect-based sentiment analysis (Section 4).

2 Distributional Semantics

The basic idea behind many modern semantic models is known as *Distributional Hypothesis*. It states that the meaning of a word is defined by the contexts where the word appears. This allows us to compare the meaning of words based on their contexts; words with similar contexts have similar meaning.

Distributional semantics models process huge amount of data in order to recognize contextual patterns. The meaning of words, phrases, or sentences is usually represented by vectors. Each word is associated with a vector, which captures all the information hidden in the contexts, including syntactic, semantic, pragmatic, or morphological information. The vectors form a k-dimensional vector space referred to as *semantic space*. The similarity of words can be measured based on their similarity (or distance) in the vector space. The most common method for measuring similarity is *cosine similarity*, which measures the cosine of the angle between the vectors of two words.

In recent years, many distributional semantics models were proposed, e.g. neural network based models (Continuous Bag-Of-Words and Skip-Gram) [6] and log-bilinear model called GloVe [7].

3 Semantic Textual Similarity

Semantic textual similarity (STS) is one of the core tasks at SemEval. Given the two textual fragments (word phrases, sentences, paragraphs, or full documents), the goal is to estimate the degree of their semantic similarity. STS systems are compared with the manually annotated data, consisting of sentence pairs and the corresponding score between 0 and 5 (higher score means higher semantic similarity). STS at SemEval 2016 were divided into English-English monolingual subtask (Section 3.1) and English-Spanish cross-lingual subtasks (Section 3.2). More information can be found in [1].

3.1 Monolingual Semantic Textual Similarity

We participated with two monolingual STS systems (the results are shown in Table 1):

- **UWB-sup**: Supervised system based on SVM regression with RBF kernel. We use state-of-the-art algorithms for the meaning representation as features. These methods benefit from various sources of information, such as lexical, syntactic, and semantic. Together, we have 301 STS features. The system is trained on all SemEval datasets from prior years (i.e. the data from 2012 up to 2015).

- **UWB-unsup**: Unsupervised system based on weighted word alignment. The method finds and aligns the words that have similar meaning and similar function in the pair of sentences.

3.2 Cross-lingual Semantic Textual Similarity

Our cross-lingual STS system for Spanish-English bilingual sentence pairs is based on two steps. Firstly, we translate Spanish sentences into English via *Google translator*. The English sentences are left untouched. Secondly, we use the same STS systems as for monolingual task. The results are shown in Table 2.

Table 1. Pearson correlations on SemEval 2016 monolingual STS evaluation data.

Team Run	Ans.-Ans.	HDL	Plagia-rism	Post editing	Ques.-Ques.	Mean	Run Rank	Team Rank
Samsung Poland	69.2	82.7	84.1	83.5	68.7	77.8	1	1
UWB-sup	62.1	81.8	82.3	82.0	70.1	75.7	2	2
UWB-unsup	64.4	79.3	82.7	81.2	53.3	72.6	21	2

Table 2. Pearson correlations on SemEval 2016 cross-lingual STS evaluation data.

Team Run	News	Multi Source	Mean	Run Rank	Team Rank
UWB-sup	90.6	81.8	86.3	1	1
UWB-unsup	91.2	80.8	86.0	2	1

4 Aspect-based Sentiment Analysis

The objective of aspect-based sentiment analysis (ABSA) is to identify the aspects of a given target entity and to estimate the sentiment polarity for each mentioned aspect. The definition of the ABSA task from SemEval 2014 distinguishes between two aspects of sentiment: aspect terms and aspect categories. The whole task is divided into four subtasks. The later SemEval ABSA tasks further distinguish between more detailed aspect categories and associate aspect terms (targets) with aspect categories.

4.1 SemEval 2014

For each subtask we propose both constrained (no external knowledge) and unconstrained approach. The constrained versions of our system are based purely on machine learning techniques. The unconstrained versions extend the constrained feature set by LDA, semantic spaces and sentiment dictionaries. The proposed approaches achieved very good results. More information can be found in [3]. Some of our results are shown in Table 3.

4.2 SemEval 2016

Our constrained submission was based on lexical and syntactic features and machine learning. The unconstrained submission additionally contained semantics features and dictionaries. We achieve state-of-the-art results in 9 experiments among the constrained systems and in 2 experiments among the unconstrained systems. We participated in four languages on both text and sentence levels. More information can be found in [2]. Some of our results are shown in Table 3.

Tab. 3. Achieved ranks and results on SemEval 2014 and 2016 ABSA task. F1 denote F_1 score in percentages.

				Constrained				**Unconstrained**			
Domain	**Year**	**Lang.**	**Level**	**Category**		**Sentiment**		**Category**		**Sentiment**	
				Rank	**F1**	**Rank**	**F1**	**Rank**	**F1**	**Rank**	**F1**
Restaurants	2016	EN	Sentence	3.	68	2.	82	8.	68	9.	82
Laptops	2016	EN	Sentence	1.	48	3.	74	7.	47	10.	74
Restaurants	2016	EN	Text	1.	81	1.	81	3.	80	1.	82
Laptops	2016	EN	Text	1.	61	1.	75	2.	60	1-2.	75
Restaurants	2014	EN	Sentence	12.	76	12.	72	7.	79	4.	78
Domain	**Year**	**Lang.**	**Level**	**Target**		**Sentiment**		**Target**		**Sentiment**	
Restaurants	2016	EN	Sentence	1.	67	4.	41	3.	67	6.	41
Restaurants	2014	EN	Sentence	1.	67	4.	41	3.	67	6.	41
Laptops	2014	EN	Sentence	5.	81	9.	73	-	-	4.	67

5 Conclusion

In this paper we presented our UWB systems for STS task at SemEval competition. We use distributional semantics models as a core part of our methods. We were ranked #2 out of 113 systems in monolingual STS and #1 out of 26 systems in cross-lingual STS. Our system for ABSA was ranked as one of the bests on both SemEval 2014 and 2016.

References

1. Brychcín, T. and Svoboda, L.: UWB at SemEval-2016 Task 1: Semantic textual similarity using lexical, syntactic, and semantic information. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016.
2. Hercig, T., Brychcín, T., Svoboda, L., and Konkol, M.: UWB at SemEval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016, pages 342-349.
3. Brychcín, T., Konkol, M., and Steinberger J.: UWB: Machine Learning Approach to Aspect-Based Sentiment Analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval-2014)*, 2014, pages 817-822.

4. Konkol, M., Brychcín, T., and Konopík, M.: Latent semantics in named entity recognition. *Expert Systems with Applications*, 2015, pages 3470-3479.
5. Brychcín, T., and Konopík, M.: Semantic spaces for improving language modeling. *Computer Speech & Language*, 2014, pages 192-209.
6. Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient estimation of word representations in vector space, 2013.
7. Pennington, J., Socher, R., and Manning, C.: Glove: Global vectors for word representation. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pages 1532-1543.