

# Sentiment Analysis in Twitter: A SemEval Perspective

Preslav Nakov

Qatar Computing Research Institute, HBKU  
Tornado Tower, floor 10  
P.O. box 5825  
Doha, Qatar  
pnakov@qf.org.qa

The recent rise of social media has greatly democratized content creation. Facebook, Twitter, Skype, Whatsapp and LiveJournal are now commonly used to share thoughts and opinions about anything in the surrounding world. This proliferation of social media content has created new opportunities to study public opinion, with Twitter being especially popular for research due to its scale, representativeness, variety of topics discussed, as well as ease of public access to its messages.

Unfortunately, research in that direction was hindered by the unavailability of suitable datasets and lexicons for system training, development and testing. While some Twitter-specific resources were developed, initially they were either small and proprietary, such as the i-sieve corpus (Kouloumpis et al., 2011), were created only for Spanish like the TASS corpus (Villena-Román et al., 2013), or relied on noisy labels obtained automatically (Mohammad, 2012; Pang et al., 2002).

This situation changed with the shared task on *Sentiment Analysis on Twitter*, which was organized at SemEval, the International Workshop on Semantic Evaluation, a semantic evaluation forum previously known as SensEval. The task ran in 2013, 2014, 2015 and 2016, attracting over 40+ of participating teams in all four editions. While the focus was on general tweets, the task also featured out-of-domain testing on SMS messages, LiveJournal messages, as well as on sarcastic tweets.

SemEval-2013 task 2 (Nakov et al., 2013) and SemEval-2014 Task 9 (Rosenthal et al., 2014) had an expression-level and a message-level polarity subtasks.

SemEval-2015 Task 10 (Rosenthal et al., 2015; Nakov et al., 2016b) further added subtasks on topic-based message polarity classification, on detecting trends towards a topic, and on determining the out-of-context (a priori) strength of association of Twitter terms with positive sentiment.

SemEval-2016 Task 4 (Nakov et al., 2016a) dropped the phrase-level subtask, and focused on sentiment with respect to a topic. It further introduced a 5-point scale, which is used for human review ratings on popular websites such as Amazon, TripAdvisor, Yelp, etc.; from a research perspective, this meant moving from classification to *ordinal regression*. Moreover, it focused on *quantification*, i.e., determining what proportion of a set of tweets on a given topic are positive/negative about it. It also featured a 5-point scale *ordinal quantification* subtask (Gao and Sebastiani, 2015).

Other related (mostly non-Twitter) tasks have explored aspect-based sentiment analysis (Pontiki et al., 2014; Pontiki et al., 2015; Pontiki et al., 2016), sentiment analysis of figurative language on Twitter (Ghosh et al., 2015), implicit event polarity (Russo et al., 2015), stance in tweets (Mohammad et al., 2016), out-of-context sentiment intensity of phrases (Kiritchenko et al., 2016), and emotion detection (Strapparava and Mihalcea, 2007). Some of these tasks featured languages other than English.

We expect the quest for more interesting formulations of the general sentiment analysis task to continue. We see SemEval as the engine of this innovation, as it not only does head-to-head comparisons, but also creates databases and tools that enable follow-up research for many years afterwards.

## References

- Wei Gao and Fabrizio Sebastiani. 2015. Tweet sentiment: From classification to quantification. In *Proceedings of the 7th International Conference on Advances in Social Network Analysis and Mining*, ASONAM '15, pages 97–104, Paris, FR.
- Aniruddha Ghosh, Guofu Li, Tony Veale, Paolo Rosso, Ekaterina Shutova, John Barnden, and Antonio Reyes. 2015. ScemEval-2015 task 11: Sentiment analysis of figurative language in Twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, SemEval '15, pages 470–478, Denver, Colorado.
- Svetlana Kiritchenko, Saif M Mohammad, and Mohammad Salameh. 2016. SemEval-2016 task 7: Determining sentiment intensity of English and Arabic phrases. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, San Diego, California.
- Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. 2011. Twitter sentiment analysis: The good the bad and the OMG! In *Proceedings of the Fifth International Conference on Weblogs and Social Media*, ICWSM '11, pages 538–541, Barcelona, Catalonia, Spain.
- Saif M Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. SemEval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, San Diego, California.
- Saif Mohammad. 2012. #Emotional tweets. In *Proceedings of \*SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task*, \*SEM '12, pages 246–255, Montreal, Canada.
- Preslav Nakov, Sara Rosenthal, Zornitsa Kozareva, Veselin Stoyanov, Alan Ritter, and Theresa Wilson. 2013. SemEval-2013 task 2: Sentiment analysis in Twitter. In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics (\*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation*, SemEval '13, pages 312–320, Atlanta, Georgia.
- Preslav Nakov, Alan Ritter, Sara Rosenthal, Veselin Stoyanov, and Fabrizio Sebastiani. 2016a. SemEval-2016 task 4: Sentiment analysis in Twitter. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, San Diego, California.
- Preslav Nakov, Sara Rosenthal, Svetlana Kiritchenko, Saif M. Mohammad, Zornitsa Kozareva, Alan Ritter, Veselin Stoyanov, and Xiaodan Zhu. 2016b. Developing a successful SemEval task in sentiment analysis of Twitter and other social media texts. *Language Resources and Evaluation*, 50(1):35–65.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '02, pages 79–86, Philadelphia, Pennsylvania.
- Maria Pontiki, Harris Papageorgiou, Dimitrios Galanis, Ion Androutsopoulos, John Pavlopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation*, SemEval '14, pages 27–35, Dublin, Ireland.
- 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 '15, pages 486–495, Denver, Colorado.
- Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2016. SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation*, SemEval '16, San Diego, California.
- Sara Rosenthal, Alan Ritter, Preslav Nakov, and Veselin Stoyanov. 2014. SemEval-2014 Task 9: Sentiment analysis in Twitter. In *Proceedings of the 8th International Workshop on Semantic Evaluation*, SemEval '14, pages 73–80, Dublin, Ireland.
- Sara Rosenthal, Preslav Nakov, Svetlana Kiritchenko, Saif Mohammad, Alan Ritter, and Veselin Stoyanov. 2015. SemEval-2015 task 10: Sentiment analysis in Twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, SemEval '15, pages 450–462, Denver, Colorado.
- Irene Russo, Tommaso Caselli, and Carlo Strapparava. 2015. SemEval-2015 task 9: CLIPeval implicit polarity of events. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, SemEval '15, pages 442–449, Denver, Colorado.
- Carlo Strapparava and Rada Mihalcea. 2007. SemEval-2007 task 14: Affective text. In *Proceedings of the International Workshop on Semantic Evaluation*, SemEval '07, pages 70–74, Prague, Czech Republic.
- Julio Villena-Román, Sara Lana-Serrano, Eugenio Martínez-Cámara, and José Carlos González Cristóbal. 2013. TASS - Workshop on Sentiment Analysis at SEPLN. *Procesamiento del Lenguaje Natural*, 50:37–44.