Injecting Sentiment Information in Context-aware Convolutional Neural Networks

Danilo Croce[1], Giuseppe Castellucci[2], and Roberto Basili[1]

[1] Department of Enterprise Engineering, University of Roma Tor Vergata, Roma Italy
{croce,basili}@info.uniroma2.it,
[2] Department of Electronic Engineering, University of Roma Tor Vergata, Roma Italy
castellucci@ing.uniroma2.it

Deep learning models achieved remarkable results in Computer Vision, Speech recognition, Natural Language Processing and Information Retrieval. In this work, we extend a Convolutional Neural Networks (CNNs) for the Sentiment Analysis in Twitter task, as this architecture achieved state-of-the-art results in [3, 4]. In particular, this architecture has been shown effective when a proper *pre-training* step is adopted to perform the early estimation of the network parameters: in [4] it is suggested to generate pre-training data starting from a randomic selection of Twitter messages annotated with simple heuristics, e.g. the presence of specific emoticons in messages. We improve the quality of such CNN architecture in two ways. First, we propose to adopt a contextual model [5] to select pre-training material from the conversations to which training messages appear, as opposed to an arbitrary selection of messages. In this way, we aim at selecting pre-training messages that could better reflect the topics of the targeted data. Second, we promote the adoption of a multi-channel schema [3] for representing the input data to the CNN. A first channel is used to accommodate lexical information provided by Distributional Models of Lexical Semantics, i.e. a vector representation of words provided by a Word Embedding. The second channel is adopted to represent sentiment oriented information as it is provided by a polarity lexicon. In particular, the sentiment oriented vectors adopted in this study refer to the automatically acquired Distributional Polarity Lexicons, as proposed in [1]. The experimental evaluation shows that the proposed solutions are beneficial w.r.t the targeted task in two languages, i.e. English and Italian. The full version of this paper is provided in [2] and it is available in the *SocialNLP@IJCAI 2016* proceedings.

References

1. Castellucci, G., Croce, D., Basili, R.: Acquiring a large scale polarity lexicon through unsupervised distributional methods. In: Proceedings of NLDB 2015 (June 2015)
2. Croce, D., Castellucci, G., Basili, R.: Injecting sentiment information in context-aware convolutional neural networks. In: Proceedings of SocialNLP@IJCAI 2016 2016 (July 2016)
3. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings EMNLP 2014. pp. 1746–1751. Association for Computational Linguistics, Doha, Qatar (October 2014)
4. Severyn, A., Moschitti, A.: Twitter sentiment analysis with deep convolutional neural networks. In: Proceedings of the SIGIR 2015. pp. 959–962. ACM, New York, NY, USA (2015)
5. Vanzo, A., Croce, D., Basili, R.: A context-based model for sentiment analysis in twitter. In: Proc. of 25th COLING: Best Paper. pp. 2345–2354. Dublin City University and Association for Computational Linguistics (2014)