Fine-grained sentiment analysis: a piece of cake? 🍰

Veronique Hoste
LT3 Language and Translation Technology Team
Department of Translation, Interpreting and Communication
Ghent University
`veronique.hoste@ugent.be`

With the emergence of the interactive Web 2.0, the amount of opinionated online text has grown immensely, as well as the interest in exploiting that information. At the same time, digitization and globalization have profoundly changed the media ecology, with an increasing trend to consume news online, more specifically via newspaper websites or through secondary gatekeepers on social media platforms, etc.

This availability of online social and curated text sources has led to a boost in sentiment analysis research, which mainly took off in early 2000 with as initial objective the identification of semantic polarity (positive, negative, or neutral) of a given text. In the last years this primary objective has evolved into a more fine-grained paradigm. This includes identifying the entity towards which a given sentiment is aimed in aspect-based sentiment analysis (Pontiki et al., 2016), identifying emotions instead of mere polarity orientations (Mohammad et al., 2018) or modeling the implicit sentiment certain events or facts convey or evoke. In this talk, I will focus on some ongoing projects in our team in which we seek to model sentiment and emotions at this fine-grained level.

Taking the aspect-based sentiment analysis framework as a starting point, I will broaden the scope from aspects to unrestricted news events and discuss our attempts to model fine-grained news events' polarity in general and economic hard news. As factual utterances often do not contain explicitly lexicalized sentiment, but rather describe "polar facts" or real-world events or objects with implied affective information, I will mainly focus on the challenges involved in modeling implicit sentiment. This implicit or prototypical polarity modeling has also been key to our work on verbal irony and the associated SemEval-2018 shared task. As irony is frequently realized through a clash between (often) explicit opinion words and a prototypically negatively connoted activity, I will discuss how we seek to model this contrast (Van Hee et al., 2018).

Finally, in view of more refined emotion modelling in text, I will elaborate on our first steps in the domain of emotion detection, starting with our pursuit of a reliable method to label emotional properties in text (De Bruyne et al., to appear). Furthermore, I will discuss the problem of choosing an appropriate framework for building an emotion-annotated corpus and our experiments on transfer learning for emotion detection in a less-resourced scenario.

De Bruyne, L., De Clercq, O. & Hoste, V. (to appear). Annotating Affective Dimensions in User-Generated Content. Comparing the reliability of best-worst scaling, pairwise comparison and rating scales for annotating valence, arousal and dominance. *Language Resources and Evaluation.*

Mohammad, S.M., Bravo-Marquez, F., Salameh, M. & Kiritchenko, S. (2018). Semeval-2018 Task 1: Affect in tweets. *Proceedings of SemEval-2018*.

Pontiki, M., Galanis, D., et al. (2016). SemEval-2016 task 5: aspect based sentiment analysis. *Proceedings of SemEval-2016*.

Van Hee, C., Lefever, E., & Hoste, V. (2018). We usually don't like going to the dentist: using common sense to detect irony on Twitter. *Computational Linguistics*, 44(4), 793–832.