

Some limits for social listening and sentiment analysis in regards to humor : an iceberg of cultural and technological subtleties behind Japanese slang and emoticons

Alice Pennors^a

^aAmstratgraph - CAE 29 Chrysalide, Brest, France

Abstract

Natural language processing and information retrieval are rapidly making progresses that contribute to sentiment analysis (Sadegh et al., 2012; Hussein, 2018), including thanks to models that rely on deep learning techniques (Kim, 2014; Lai et al., 2015; Chen et al., 2017; Lin et al., 2017). Analyzing Japanese, however, has long raised challenges for automated analysis (Pardamean et al., 2019), mainly due to the sparse and highly contextual nature of the language itself, which uses no whitespaces, ambiguous sentences, unpredictable morphology and grammar features that diverge from most European languages, not to mention multiple ways to split characters into words, transcriptions and abbreviations, as well as the spontaneous use of many colloquial elements, such as emoticons, filler expressions and local idiomatizations (Imamura et al., 2008).

This already induces completely different possible readings, meanings and interpretations that require from trained translators or analysts to know how to "read between the lines", while bringing many occasions for puns and play of words that can be hard to grasp. It similarly complexifies the tasks that automated tools may perform to analyze and eventually predict syntactic structure, dependency parsing and sequential tagging (Bataa and Wu, 2019), even more so given that Japanese language is far from being static. On the contrary, it is very alive and dynamic in developing trendy neologisms and lexical stocks influenced by foreign, modern, and digital cultures (Ibatullina et al., 2019).

As far as I could observe in my professional work as strategic analyst for the past 15 years, social listening platforms may well be relying on increasingly accurate automated semantic analysis tools, which prove very useful to embrace thousands online publications "in one click" for opinion mining, sentiment still has to rely on culturally relevant, linguistically skilled human readers. In my experience, this is especially true outside of Western cultures, the Japanese Web being my own specific play ground. In this paper, I shall briefly illustrate some technological and cultural complexities that lie behind the assessment of written humor and irony through a few examples among the most popular slang abbreviations and emoticons used by Japanese net users and their own evolutions.

Keywords

Social listening, Sentiment analysis, Humor, Emoticons

Words/Machines-2021: Do machines have a sense of humor?, March 5, 2021, University of Western Brittany, Faculty of Arts, Brest, France

✉ contact@amstratgraph.eu (A. Pennors)

🌐 <http://amstratgraph.eu> (A. Pennors)



© 2021 Use permitted under Creative Commons License Attribution ShareAlike 4.0 International (CC BY SA 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)