

Identifying product failures from reviews in noisy data by distant supervision

Tutubalina E.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© Springer International Publishing Switzerland 2016. Product reviews contain valuable information regarding customer satisfaction with products. Analysis of a large number of user requirements attracts interest of researchers. We present a comparative study of distantly supervised methods for extraction of user complaints from product reviews. We investigate the use of noisy labeled data for training classifiers and extracting scores for an automatically created lexicon to extract features. Several methods for label assignment were evaluated including keywords, syntactic patterns, and weakly supervised topic models. Experimental results using two real-world review datasets about automobiles and mobile applications show that distantly supervised classifiers outperform strong baselines.

http://dx.doi.org/10.1007/978-3-319-45880-9_12

Keywords

Distant supervision, Opinion mining, Problem phrase extraction, Product defects