

Applied Intelligent Data and Text Analytics

Sentiment Labelled Sentences Dataset

Dataset Information:

This dataset was created for the following:

Kotzias, D., Denil, M., Freitas, N.D., & Smyth, P. (2015). From Group to Individual Labels Using Deep Features. *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

It contains sentences labelled with positive or negative sentiment.

Format:

Sentence score.

Details:

Score is either 1 (for positive) or 0 (for negative).

The sentences come from three different websites/fields:

- imdb.com
- amazon.com
- yelp.com

For each website, there exist 500 positive and 500 negative sentences. Those were selected randomly for larger datasets of reviews.

The authors attempted to select sentences that have a clearly positive or negative connotation; the goal was for no neutral sentences to be selected.

Attribute Information:

The attributes are text sentences, extracted from reviews of products, movies, and restaurants.

Link:

<https://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences>