

UCI



Machine Learning Repository

[Center for Machine Learning and Intelligent Systems](#)

[About](#) [Citation Policy](#) [Donate a Data Set](#) [Contact](#)

☒ Repository ☐ Web

Google™

[View ALL Data Sets](#)

Drug Review Dataset (Drugs.com) Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: The dataset provides patient reviews on specific drugs along with related conditions and a 10 star patient rating reflecting overall patient satisfaction.

Data Set Characteristics:	Multivariate, Text	Number of Instances:	215063	Area:	Life
Attribute Characteristics:	Integer	Number of Attributes:	6	Date Donated	2018-10-04
Associated Tasks:	Classification, Regression, Clustering	Missing Values?	N/A	Number of Web Hits:	37722

Source:

Surya Kallumadi
Kansas State University
Manhattan, Kansas, USA
surya '@' ksu.edu

Felix Gräßer
Institut für Biomedizinische Technik
Technische Universität Dresden
Dresden, Germany
felix.graesser '@' tu-dresden.de

Data Set Information:

The dataset provides patient reviews on specific drugs along with related conditions and a 10 star patient rating reflecting overall patient satisfaction. The data was obtained by crawling online pharmaceutical review sites. The intention was to study

- (1) sentiment analysis of drug experience over multiple facets, i.e. sentiments learned on specific aspects such as effectiveness and side effects,
- (2) the transferability of models among domains, i.e. conditions, and
- (3) the transferability of models among different data sources (see 'Drug Review Dataset (Druglib.com)').

The data is split into a train (75%) a test (25%) partition (see publication) and stored in two .tsv (tab-separated-values) files, respectively.

Important notes:

When using this dataset, you agree that you

- 1) only use the data for research purposes
- 2) don't use the data for any commercial purposes
- 3) don't distribute the data to anyone else
- 4) cite us

Attribute Information:

1. drugName (categorical): name of drug
2. condition (categorical): name of condition
3. review (text): patient review
4. rating (numerical): 10 star patient rating
5. date (date): date of review entry
6. usefulCount (numerical): number of users who found review useful

Relevant Papers:

Felix Gräßer, Surya Kallumadi, Hagen Malberg, and Sebastian Zaunseder. 2018. Aspect-Based Sentiment Analysis of Drug Reviews Applying Cross-Domain and Cross-Data Learning. In Proceedings of the 2018 International Conference on Digital Health (DH '18). ACM, New York, NY, USA, 121-125. DOI: [\[Web Link\]](#)

Citation Request:

Felix Gräßer, Surya Kallumadi, Hagen Malberg, and Sebastian Zaunseder. 2018. Aspect-Based Sentiment Analysis of Drug Reviews Applying Cross-Domain and Cross-Data Learning. In Proceedings of the 2018 International Conference on Digital Health (DH '18). ACM, New York, NY, USA, 121-125. DOI: [\[Web Link\]](#)

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)