Dataset

Sentiment140 dataset with 1.6 million tweets

Sentiment analysis with tweets

Μαριος Μιχαηλιδης KazAnova • updated 3 years ago (Version 2)

Data Tasks Notebooks (128) Discussion (10) Activity Metadata

Download (228 MB) New Notebook

Your Dataset download has started.
Show your appreciation with an upvote

812

🖤 Usability 8.8	⚖ License Other (specified in description)	🏷 Tags internet, online communities, social networks, linguistics, languages

Description

Context

This is the sentiment140 dataset. It contains 1,600,000 tweets extracted using the twitter api . The tweets have been annotated (0 = negative, 4 = positive) and they

---

Data Explorer

227.74 MB

🖽 training.1600000.processed...

Summary

▸ 📁 1 file

▸ 🖽 6 columns

← **training.1600000.processed.noemoticon.csv** (227.74 MB)   ⬇ 🔲 ⛶
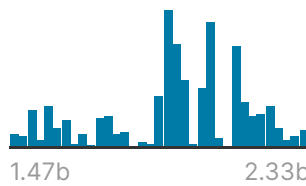
Detail Compact Column   6 of 6 columns ⌄

About this file

This is the sentiment140 dataset.
It contains 1,600,000 tweets extracted using the twitter api . The tweets have been annotated (0 = negative, 2 = neutral, 4 = positive) and they can be used to detect sentiment .
It contains the following 6 fields:

1. **target**: the polarity of the tweet (*0* = negative, *2* = neutral, *4* = positive)

# 0 target	# 1467810369 id	A Mon Apr 06 22:1... date	A NO_QUERY flag	A _TheSpecialOne_ user	A te... te
0 — 4	1.47b — 2.33b	774362 unique values	1 unique value	659775 unique values	
0	1467810672	Mon Apr 06 22:19:49 PDT 2009	NO_QUERY	scotthamilton	is ca Fa it as tc
0	1467810917	Mon Apr 06 22:19:53 PDT 2009	NO_QUERY	mattycus	@k ma ba sa ou
0	1467811184	Mon Apr 06 22:19:57 PDT 2009	NO_QUERY	ElleCTF	my it or
0	1467811193	Mon Apr 06 22:19:57 PDT 2009	NO_QUERY	Karoli	@r al i ca o.
0	1467811372	Mon Apr 06 22:20:00 PDT 2009	NO_QUERY	joy_wolf	@k wh
0	1467811592	Mon Apr 06 22:20:03 PDT 2009	NO_QUERY	mybirch	Ne
0	1467811594	Mon Apr 06 22:20:03 PDT 2009	NO_QUERY	coZZ	@L ti Ra bi th
0	1467811795	Mon Apr 06 22:20:05 PDT 2009	NO_QUERY	2Hood4Hollywood	@1 di
0	1467812025	Mon Apr 06 22:20:09 PDT 2009	NO_QUERY	mimismo	@t mu
0	1467812416	Mon Apr 06 22:20:16 PDT 2009	NO_QUERY	erinx3leannexo	sp pl sr
0	1467812579	Mon Apr 06 22:20:17 PDT 2009	NO_QUERY	pardonlauren	I ea
0	1467812723	Mon Apr 06 22:20:19 PDT 2009	NO_QUERY	TLeC	@c co