README

The file `twitter-immigration-debates.csv` provides the tweets involved into the debates about the italian immigration policies, collected over 1-month period from January 2019, 23rd to February 2019, 22nd. The botometer's scores are collected over January 2020.

Statistics

- 243,226 tweets
- 796,456 retweets
- 97,735 users
- 573 verified accounts
- 87,487 human-operated accounts
- 3,263 automated accounts
- 6,985 unclassified accounts

Data Format

The data is in comma-separated format with the following headers:

> user_id, tweet_id, retweet, timestamp, user_verified, cap_universal, scores_universal, bot

ATTRIBUTE	TYPE	Description
`user_id`	Int64	The integer representation of the unique identifier for the User.
`tweet_id`	Int64	The integer representation of the unique identifier for this Tweet.
`retweet`	Boolean	The boolean indicates if the record is a tweet or a retweet
`timestamp`	Timestamp	Timestamp of the tweet. Example: `Mon Jun 11 19:03:01 +0000 2018`
`user_verified`	Boolean	When true, indicates that the user has a verified account.
`cap_universal`*	Real	Universal Complete Automation Probability (CAP) by Botometer.
`scores_universal`*	Real	Universal scores by Botometer.
`bot`*	Boolean	When true, indicates that the user is classified as bot. `Note. The classification consider CAP universal score and 0.4 as a threshold.`

* When the value is `?`, the account results as unclassified.