

EURO 2020 Italian Tweet Dataset Documentation

December 17, 2023

This document provides documentation for the EURO 2020 tweet dataset. The dataset includes anonymized tweets extracted during the EURO 2020 event. Details about the structure of each tweet, hashtag classifications, and user stances are explained below.

Dataset Description

The dataset is comprised of the following files:

- `EURO2020_italian_tweets.json`: Anonymized tweets during EURO 2020.
- `hashtags_classification.json`: Hashtags with associated classifications.
- `users_stances.json`: Users' stances based on tweet hashtags.

Tweet Structure

Each tweet contained in the `EURO2020_italian_tweets.json` file follows the structure outlined below:

Listing 1: Tweet Structure

```
{
  "tweet_id": Unique identifier for each tweet (str),
  "user": Anonymized author id (str),
  "date": Date and time of the tweet in ISO 8601 format (str),
  "hashtags": List of hashtags used in the tweet (list),
  "mentions": Mentioned users in the tweet (list),
  "retweets": List of retweets for the tweet (list),
  "reply_to": Users to whom the tweet is a reply (list),
  "quote_to": Users whose tweet is quoted (list),
  "followers": Number of followers of the tweet author (int),
  "following": Number of accounts the tweet author is following (int),
  "tweets": Total number of tweets by the author (int),
  "tweet_classification": Tweet stance according to hashtags classification (int),
  "user_classification": User stance according to hashtags classification (float),
}
```

Hashtag Classification

The `hashtags_classification.json` file contains hashtags with associated numerical classifications. The classifications are as follows:

- ± 3 : Strongly against (+)/strongly in favor(-) of the act of kneeling.
- ± 1 : Significant but not specific on the issue of "kneeling."
- 0: Neutral or irrelevant hashtags.

User Stances

The `users_stances.json` file contains users' stances based on the hashtags present in their tweets. Stances are computed according to the associated hashtag classifications.