Readme for Facebook Likers Dataset

Updated on September 14, 2016

1. Data Description

We provide Facebook Likers dataset collected in 2015. The dataset was used in a paper “Uncovering Fake Likers in Online Social Networks” in CIKM 2016. It contains 16 feature values extracted from 6,895 fake likers and 6252 legitimate likers.

2. Paper Citation

Please cite the following paper when using the dataset.


3. Data Format

“all_likers.csv” contains 16 feature values of 6,895 fake likers and 6252 legitimate likers in the form of “total # of posts created by user, average # of URLs per post, proportion of shared photos out of total # of posts, proportion of shared pages out of total # of posts, proportion of shared posts out of total # of posts, average # of posts per day, maximum # of posts in a day, skewness of daily posted posts, category entropy, proportion of verified pages out of total # of pages liked by a user, average # of likes per post, average # of comments per post, average # of shares per post, longevity, # of lines in About section, # of friends, class”.

Detailed information regarding each feature is described in the CIKM paper.

4. Contact

Please feel free to contact Kyumin Lee if you have any question about the dataset.

Email: kyumin.lee AT usu DOT edu

5. Agreement

The Facebook Likers dataset is strictly prohibited to de-anonymize Facebook users from the dataset.

The dataset is licensed under a Creative Commons Attribution-Noncommercial 3.0 United States License.