Scott Breitbach

09-November-2021

DSC630 Predictive Analytics

Assignment 11.2: Case Study

# Online Privacy Foundation: Predicting Psychopathy on Twitter

## Overview

In 2011 a Cornell research study was released showing a link between language characteristics used and psychopathy. Writings from known murderers was assessed and patterns were found that could indicate likelihood of psychopathy. Following this paper, the media began to ask about the implications of this study. Could one determine psychopathy from user’s social media? Could law enforcement use this to track down potential psychopathic killers?

These questions sparked the interest of researchers with the Online Privacy Foundation, a volunteer organization “whose mission is to empower individuals to make informed choices about how they use social media and interact online” (onlineprivacyfoundation.org). The Online Privacy Foundation partnered with Florida Atlantic University and Kaggle in a predictive analytics competition to “predict people with a sufficiently high degree of Psychopathy based on Twitter usage and Linguistic Inquiry” (Kaggle).

## Data Understanding

In order to determine psychopathy, over 3,000 Twitter users volunteered to take a personality test and offer up their twitter history for analysis. The self-assessment consisted of 40 questions, rating users from 1-5 on 8 personality traits, consisting of the Big 5 (openness, conscientiousness, extraversion, agreeableness, and neuroticism) and the ‘dark triad’ (psychopathy, Machiavellianism, and narcissism). Determining these personality traits was based on a checklist developed by Professor Del Paulhus at the University of British Colombia, as well as the Hare Psychopathy Checklist (PCL-R).

The Online Privacy Foundation already had some experience in this field as they had previously run a “Facebook Big 5” experiment, comparing users’ Facebook activity to their personality profile. They anticipated that Twitter would be easier because the Twitter data is more public than Facebook. Ultimately, they were able to obtain an average of over 3,000 tweets per person.

## Data Preparation

As one would expect from an organization that promotes online privacy, the data was anonymized prior to being provided for the competition. In addition, rather than providing the Twitter content, the Twitter data itself was converted into 337 numerical features so that no identifiable information remained when it was made public. The types of information extracted into features were generally metadata, like a user’s total number of tweets, how frequently they tweet, the number of followers they have, and their Klout score (which is a measure of the user’s influence on social media).

In addition to this metadata, linguistic analysis was performed, using numbers to represent textual information. Traits that correlated with undesirable qualities included things like curse words, angry responses, words like “hate” and “we”, use of periods, and filler words like “blah”, “I mean”, and “you know”.

## Modeling

Of the initial 3,000+ volunteers, the data was compiled from 2,927 participants. The goal was to identify high psychopathy scores, which was defined as greater than 2 standard deviations above the mean value of 1.98. This target group comprised only about 3% of the data set, making it very imbalanced for modeling. Of these, about 41 people were considered “certifiable,” meaning they scored highly on all three attributes of the dark triad. This was the exception, however, and most people fell on a spectrum for each quality.

The entire data set was split in half into training and testing sets for the competition, with the target variable (individual user’s personality trait scores) removed from the test set.

## Deployment

While researcher Chris Sumner had hoped to disprove a connection between one’s Twitter feed and psychopathy traits, the competition proved to be very successful. “With machine learning you can really increase the odds of making an accurate prediction,” says Sumner. “It far exceeded my expectations.”

One of the biggest concerns with this tool is the potential for false positives. There are innocent people who may score highly on some or many of these darker traits but are perfectly law-abiding citizens and productive members of society. As a case in point, many of the best CEOs tend to score highly on the psychopathy trait.

## Summary and Conclusions

The positive findings of this study raise societal concerns over potential uses of this tool, either by individuals or organizations, for good or for bad. The authors raised concerns that use of such tools against individuals would impact people use and enjoyment of the Internet, though it could also be useful to provide a look at societal trends over time.

Employers are already using individuals’ social media content to pre-screen for employment, which can be problematic. Can you really tell if a person will be a good employee based on their social media presence? While you may be able to find some real red flags, there may be bias based on differences of opinion that would have no bearing on whether or not that person would make a “good employee.” Similarly, predictive algorithms are being used to determine the likelihood that a juvenile will reoffend. These sort of examples open the discussion of “what are the acceptable uses for this stuff and what’s unacceptable?”

## Sources

Hancock, Jeffrey T., et al. “Hungry like the Wolf: A Word‐Pattern Analysis of the Language of Psychopaths.” British Psychological Society, John Wiley &amp; Sons, Ltd, 14 Sept. 2011, <https://bpspsychub.onlinelibrary.wiley.com/doi/full/10.1111/j.2044-8333.2011.02025.x>.

Hill, Kashmir. “Using Twitter to Identify Psychopaths.” Forbes, Forbes Magazine, 29 July 2012, <https://www.forbes.com/sites/kashmirhill/2012/07/20/using-twitter-to-help-expose-psychopaths/?sh=4edac70d8463>.

The Online Privacy Foundation, <https://onlineprivacyfoundation.org/>.

“Psychopathy Prediction Based on Twitter Usage.” Kaggle, 2011, <https://www.kaggle.com/c/twitter-psychopathy-prediction/overview/description>.

Sumner, Chris, et al. “Predicting Dark Triad Personality Traits from Twitter Usage and a Linguistic Analysis of Tweets.” 2012 11th International Conference on Machine Learning and Applications, 2012, <https://doi.org/10.1109/icmla.2012.218>.

Sumner, Chris. DEF CON® 20 Hacking Conference - Speakers. 2012, <https://defcon.org/html/defcon-20/dc-20-speakers.html#Sumner>.