Profiling Users Based on Tweets

Aim: To profile users (especially politicians) based on social media activity to help the voters decide whom to vote.

Our generation especially young people spend most of their time on smartphones. We even get our news from social media and most of us don’t even bother questioning the content of the news. We don’t vote as much as we should. (as can be seen from the statistics of Brexit or any other election). The reason behind this is that we don’t want to spend energy on investigating every single candidate and fact-check every single article to make a decision about whom to vote. The term “FAKE NEWS” and “Post-Truth” etc were coined recently only. We’re overwhelmed with too much information (true or fake) and can’t tell what’s what anymore. Therefore, fact checking websites and apps are being developed to keep the fakeness in check.

I believe that a fully automated highly accurate fact check algorithm will take years to develop. So what can we do to help people decide swiftly whom to vote when they can’t trust the politician’s words or news? The answer is using the power of **data science to profile a user based on his social media activity** (Twitter in this case). More and more people are using tweets as a way of announcing, denouncing, trolling, spreading rumors, truths, lies etc. Now this provides us a very personal dataset to look into the mind of a user and figure out his personality. Also, adding fact-check to this will give us a statistics of how frequently a person lies/misinforms or spread fake news. This information can then be presented on a single card to summarize the personal traits of a user.

Applying this to politicians will give the users enough information about the character of the candidate to make decisions on whom to vote. Turning this idea into an app will encourage even younger people to vote as they’ll have all the knowledge about a candidate summarized, which takes out their effort in researching every single candidate and make it easier to choose and hence vote.

I have done an analysis of 3200 tweets of POTUS Donald Trump to show how this method can work. I found out that he uses 2 different phones to post tweets and most of his angry, emotional tweets come from an android phone. This shows that his personal phone is the android one. Most of the official news, announcements, hash tags come from iPhone so maybe that’s the official use phone. And as the data suggested that since Trump became POTUS the tweets from the personal phone increased for a month but then exponentially decreased and now it’s totally zero implying he’s not using the android phone anymore. Now this means that he changed his phone from the end of March 2017. This means he either changed it by his own will or the phone broke down or he was forced to change the phone because of his bold tweets that got him into controversial news again and again. He’s been tweeting 80 tweets a month on average from last year till Feb 2017 but then he hit his all time low of approx.. 30 tweets in March. He didn’t stop tweeting suddenly but asymptotically went o zero which suggests that his officials more or less forced him to keep his tweets in check.

Looking at the ratio of adjectives posted between iphone and andoid most of the negative adjectives like dishonest, crazy, terrible, fake etc are likely to came from android and topics like world bank, terrorist,#makeamericagreatagain etc came from iPhone.

Taking a look at the Politifact apps findings we can see that his tweets were false news quite a lot of time. Combining these two information we can safely hypothesize that Donald trump is a difficult to deal with person with low control on anger and who can lie about data to make his point.

His most retweeted tweet was when he attacked media by calling them dishonest and he attacked media again and again which shows he loves to rile up the people with what they want to hear without any proof.

Based on this data we can calculate an index that will summarizes the personality and that can be used to compare politicians/users.