The use of sentiment analysis and methods like LSD have become increasingly popular with analyzing political texts. However, they are not always optimal at determining policy stances. For instance, even though positive sentiment may indicate being pro-gay marriage if a tweet says “I support gay marriage,” more nuance comes in when policies are less black and white, or if positive/negative sentiment does not line up well with a certain position—for instance, if text is anti-Brexit, it might have a negative sentiment attached to it, even though it is pro-European.

The authors propose the question: “Is it reasonable to assume that the sentiment of a document reflects the stance of that document toward the primary topic of the document?”

The authors try to take this, sentiment-stance correlation, and measurement bias that is associated with it into account by looking at three examples:

1. Sentiment and Stance in Tweets About the 2017 Women’s March

* Using VADER, they analyzed sentiment of 2.5 million tweets about the march. General sentiment was positive
* Then, 20,000 Women’s March tweets were “labeled them by hand according to whether they used generally positive or generally negative language, and whether they indicated approval or opposition for the movement”
* Authors found that sentiment and stance were weakly correlated—that is, women in support of the march may have tweeted with negative sentiment, and vice versa.
* Next, the authors trained a BERT model to classify tweets as either pro or anti march. They then applied this to the replication data tweets, and compared the BERT stance prediction results with the tweets respective sentiment, calculated by VADER.
* Finally, the authors look at another random sample of 1500 tweets that were coded to show ideology on a scale of -2.5 to 2.5 via Bayesian Ideal Point Estimation (Barbera). They then used a dichotomous VADER classifier (positive or negative) as well as their BERT classifier (left vs right) on the same tweets, and regressed the results. While they find a strong association between liberal ideology and approval of the movement, the coefficient for ideology in the VADER model is substantially smaller than the equivalent coefficient in the BERT model

1. Sentiment and Stance in Open-Ended Survey Responses About Donald Trump

* Moon of the Nation (MOTN) conducted surveys from November 2016 to September 2019 about the state of US politics. A corpus of 7,146 responses that mentioned Trump was collected.
* First, authors compared the true sentiment of each response to sentiment predictions from LSD and VADER. They also used ML models. Generally, ML models performed better than dictionaries.
* Then, using the model from the first example, authors get a sample of the MOTN corpus and generate dichotomous labels showing predicted sentiment and stance. Logistics models are then fitted to see if these labels can predict actual ideology of survey respondents (i.e. what they said in the survey—if they support Trump or not).
* Generally, those who are more conservative were morel likely to support Donald trump. “As in the Women’s March example, both models using machine-generated measures capture the expected direction of the relationship, but in the sentiment-trained model, Trump approval among liberals is being significantly overestimated compared to the stance-trained model.”

1. Sentiment and Stance in Tweets About the Kavanaugh Confirmation

* Unlike the previous two examples where sentiment can serve as a proxy for stance, here the authors mix things up by offering an example where sentiment may be the opposite of stance—the nomination of Brett Kavanaugh to the Supreme Court in 2018. That is, people supporting Christine Blaisey Ford who testified against Kavanaugh may have positive sentiments in their tweets, but they are against Kavanaugh’s nomination. Likewise, supporters of Kavanaugh may have negative sentiment tweets if their tweets illustrate frustration with the nomination process.
* Authors look at a corpus of tweets about the Kavanaugh nomination.