Using tweets to predict real-time food shortages

April 12 2022

Credit: Marten Bjork via Unsplash

The sentiments and emotions expressed in tweets on Twitter can be used

in real time to assess where supply chain disruptions due to a pandemic, war or natural disaster may lead to food shortages, according to researchers at Penn State and the Hamad Bin Khalifa University in Qatar. They found that food security-related tweets that expressed anger, disgust or fear were strongly correlated with actual food insufficiency in certain U.S. states early in the COVID-19 pandemic. These findings can potentially be used to develop a low-cost early warning system for identifying where food-security interventions are most needed, according to the researchers.

"The onset of the COVID-19 pandemic and its related supply-chain disruptions prompted worldwide concerns about food access and availability, and many people took to social media to express these concerns," said Stephan Goetz, professor of agricultural and regional economics at Penn State and director of the Northeast Regional Center for Rural Development (NERCRD). "We wanted to see if real-time tweets could be used to identify specific states or regions facing food supply or insecurity problems."

Goetz said that rather than looking only at the number of tweets related to food insufficiency, he and his colleagues wanted to know how people actually felt about their food situation. Using artificial intelligence, they identified the sentiments and emotions associated with the tweets, which allowed them to separate tweets expressing concerns about the food supply from those expressing relief or contentment.

In the tweets analyzed, the most commonly expressed emotion was joy, which the researchers interpreted as a reflection of widespread relief that the U.S. food supply remained relatively stable during the pandemic, despite early concerns about food shortages. Tweets that expressed negative emotions such as anger, disgust and fear were strongly correlated with actual food insufficiency problems in certain states.

To carry out the study, which was published online last week in the journal *Applied Economic Perspectives and Policy*, the researchers used a dataset known as GeoCoV19, which contains hundreds of millions of COVID-19-related tweets in multiple languages from all over the world. They selected only English language tweets that originated in the U.S. and were posted between Feb. 1, 2020, and Aug. 31. They further narrowed the resulting set of tweets to include only those that contained language related to food insecurity.

"We curated a list of 138 words and phrases—such as 'can't afford groceries,' or 'food prices' or 'food shortage'—which we used to isolate the tweets most relevant to our research question, and ended up with more than 1.2 million tweets in our data set," said Connor Heaton, doctoral candidate in Penn State's College of Information Sciences and Technology and co-author of the study. "Then, we used what are known as contextual language models to extract the sentiment and emotion based in each tweet. These models leverage the emergent meaning in a sequence of words rather than relying on a predefined relation between individual words and a sentiment and/or emotion."

The artificial intelligence language model they used can detect sentiments of negative, neutral, and positive; and emotions of anger, disgust, fear, joy, sadness, surprise or neutral.

For example, this tweet was classified as having positive sentiment and expressing joy: "God bless America's food supply chain, from the producers to the distributors to the grocers: The food supply chain, they say, remains intact and has been ramping up to meet the unprecedented stockpiling brought on by the coronavirus pandemic."

In contrast, this tweet was classified as negative and expressing fear: "A food bank executive in Louisiana who worked through Hurricane Katrina said he had never witnessed such a combination of need, scarcity

and anxiety. 'Crazy' pretty much sums it up,' he said."

The researchers also mapped the tweets at the U.S. state level so they could compare their findings against actual food insufficiency data from the U.S. Census Household Pulse Survey (HPS), a weekly national survey administered online to a randomly selected representative sample of U.S. households, which asks participants whether they sometimes or often did not have enough food, both in the previous week and prior to the onset of the pandemic.

When averaged across all states over the six-month period, tweets expressing the emotions of anger, disgust and fear were significantly correlated with actual state-level food insufficiency rates reported in the HPS. At the state level, tweets that expressed fear were most strongly correlated with actual food insufficiency in California, Illinois, New York, Texas and Wisconsin. Tweets expressing anger or disgust were positively correlated with food insufficiency in 12 states.

Goetz said that further research is needed to assess whether this method of identifying localized food-security emergencies can replace more expensive and time-consuming methods, such as surveys.

More information: Stephan J. Goetz et al, Food insufficiency and Twitter emotions during a pandemic, *Applied Economic Perspectives and Policy* (2022). DOI: 10.1002/aepp.13258

Provided by Pennsylvania State University