Want to know how processed your food is? There's an algorithm for that

June 2 2023, by Tanner Stening

Credit: Pixabay/CC0 Public Domain

Northeastern researchers have been busy trying to better understand the links between "ultra-processed foods" and human health through the university-sponsored Foodome project.

As part of that effort, researchers with the the Center for Complex Network Research have now developed a machine learning algorithm they say accurately predicts the degree of processing in [food products](#) that make up the U.S. food supply. Their findings were published in *Nature Communications* in April.

The machine learning classifier, called FoodProX, uses nutritional labeling information provided by the U.S. Department of Agriculture's Food and Nutrient Database for Dietary Studies as inputs to score the level of processing in a given food product.

The algorithm works by producing an output that represents the likelihood that a respective food falls into one of the four categories that are part of the NOVA food classification system—a system developed by researchers at the University of São Paulo, Brazil, that the researchers say is "widely used in [epidemiological studies](#)."

Users can try out the tool by visiting TrueFood research project's website. Users can search for a food to see its food processing score. The algorithm assigns each product a single score between zero (which denotes "minimally or unprocessed" food) and 100 (highly ultra-processed food).

Using FoodProX, researchers were able to bridge gaps in the Nutrient Database for Dietary Studies; classify "complex recipes and mixed foods and meals"; and provide a higher resolution lens with which to examine processed foods. As a result, the researchers note that FoodProX provides a sharper understanding of just how processed foods actually are—an important step for researchers studying the health impacts these foods have.

The researchers note how the NOVA system, which splits foods into four classifications, from "unprocessed or minimally processed" to ultra-

processed, is fundamentally limiting because it doesn't account for the different gradations of processing within each separate category.

"These perceived homogeneity of NOVA 4 foods limits both scientific research and practical consumer guidance on the health effects of differing degrees of processing," the researchers wrote. "It also reduces the industry's incentives to reformulate foods towards less processed offerings, shifting investments from the ultra-processed NOVA 4 foods to the less processed NOVA 1 and NOVA 3 categories."

"In the paper what we do is really say that we believe that nutritional information, so the chemicals that are measured as nutrients in the nutritional facts, somehow encode the fingerprint of food processing," says Giulia Menichetti, senior research scientist at Northeastern's Network Science Institute and lead author of the research. "Because when we process a food, when we modify some staple ingredients, we change its chemistry in many different ways."

That "fingerprinting" is the way that researchers can glean insight into just how many chemical alterations have been made to a given food.

"We don't necessarily know what all the chemical fingerprints are that are associated, one-to-one, with each process," Menichetti told Northeastern Global News. "We can't even enumerate how many different ways there are to process a food."

Ultimately, the AI tool confirmed the team's prior finding that more than 73% of the U.S. food system is ultra-processed, while providing a level of detail not previously obtainable. Menichetti says her team is the first to successfully create an AI tool that reliably assesses the chemical content of food.

"It's the first paper in the space of nutrition and public health that

leverages machine learning to reproducibly and systematically score foods according to their degree of food processing," she says.

The team's work is significant because, as Menichetti says, "there wasn't much of a data culture" in the field of nutrition and health science as it relates to food processing, which promoted less scientifically rigorous conversations about what processing even means.

"When you don't have a systematic way to look at a [food](#) and assess its properties, then it's hard to enable large studies elsewhere in the world that are comparable," Menichetti says.

"FPro helps us assess an individual's diet quality, offering predictive power over 200+ health variables," says Albert-László Barabási, Robert Gray Dodge Professor of Network Science at Northeastern and co-author of the study. "It tells us the impact of replacing processed foods with less processed alternatives of the same item, resulting in personalized dietary shifts with minimal effort."

Provided by Northeastern University