

A new way to help computers recognize patterns

January 25 2006

Researchers at Ohio State University have found a way to boost the development of pattern recognition software by taking a different approach from that used by most experts in the field. This work may impact research in areas as diverse as genetics, economics, climate modeling, and neuroscience.

Alex Martinez, assistant professor of electrical and computer engineering at Ohio State, explained what all these areas of research have in common: pattern recognition.

He designs computer algorithms to replicate human vision, so he studies the patterns in shape and color that help us recognize objects, from apples to friendly faces. But much of today's research in other areas comes down to finding patterns in data -- identifying the common factors among people who develop a certain disease, for example.

In fact, the majority of pattern recognition algorithms in science and engineering today are derived from the same basic equation and employ the same methods, collectively called linear feature extraction, Martinez said.

But the typical methods don't always give researchers the answers they want. That's why Martinez has developed a fast and easy test to find out in advance which algorithms are best in a particular circumstance.

"You can spend hours or weeks exploring a particular method, just to

find out that it doesn't work," he said. "Or you could use our test and find out right away if you shouldn't waste your time with a particular approach."

The research grew out of the frustration that Martinez and his colleagues felt in the university's Computational Biology and Cognitive Science Laboratory, when linear algorithms worked well in some applications, but not others.

In the journal *IEEE Transactions on Pattern Analysis and Machine Intelligence*, he and doctoral student Manil Zhu described the test they developed, which rates how well a particular pattern recognition algorithm will work for a given application.

Along the way, they discovered what happens to scientific data when researchers use a less-than-ideal algorithm: They don't necessarily get the wrong answer, but they do get unnecessary information along with the answer, which adds to the problem.

He gave an example.

"Let's say you are trying to understand why some patients have a disease. And you have certain variables, which could be the type of food they eat, what they drink, amount of exercise they take, and where they live. And you want to find out which variables are most important to their developing that disease. You may run an algorithm and find that two variables -- say, the amount of exercise and where they live -- most influence whether they get the disease. But it may turn out that one of those variables is not necessary. So your answer isn't totally wrong, but a smaller set of variables would have worked better," he said. "The problem is that such errors may contribute to the incorrect classification of future observations."

Martinez and Zhu tested machine vision algorithms using two databases, one of objects such as apples and pears, and another database of faces with different expressions. The two tasks -- sorting objects and identifying expressions -- are sufficiently different that an algorithm could potentially be good at doing one but not at the other.

The test rates algorithms on a scale from zero to one. The closer the score is to zero, the better the algorithm.

The test worked: An algorithm that received a score of 0.2 for sorting faces was right 98 percent of the time. That same algorithm scored 0.34 for sorting objects, and was right only 70 percent of the time when performing that task. Another algorithm scored 0.68 and sorted objects correctly only 33 percent of the time.

"So a score like 0.68 means 'don't waste your time,'" Martinez said. "You don't have to go to the trouble to run it and find out that it's wrong two-thirds of the time."

He hopes that researchers across a broad range of disciplines will try out this new test. His team has already started using it to optimize the algorithms they use to study language and cancer genetics.

Source: Ohio State University

Citation: A new way to help computers recognize patterns (2006, January 25) retrieved 1 May 2024 from <https://phys.org/news/2006-01-patterns.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--