

New machine learning method predicts future data patterns to optimize data storage

February 15 2024



Credit: CC0 Public Domain

Researchers have developed a new machine-learning technique that helps computer systems predict future data patterns and optimize how

information gets stored. They found these predictions could provide up to a 40% speed boost on real-world data sets.

In a [paper](#), posted to the *arXiv* preprint server and presented as a spotlight at the [Conference on Neural Information Processing Systems \(NeurIPS\)](#) in December 2023, researchers from Carnegie Mellon University and Williams College shared that this new method could lead to significantly faster databases and more efficient data centers.

They discussed a common data structure called a list labeling array, which stores information in sorted order inside a computer's memory. Keeping data sorted allows a computer to find it quickly, like how alphabetizing a long list of names makes it easy to locate someone.

However, efficiently maintaining the sorted order as new data comes in can be challenging. Until now, computer systems could only prepare for the [worst-case scenario](#), constantly moving data around to make room for new items. This can be slow and computationally expensive.

This new machine learning method gives these data structures the power to predict. The computer analyzes patterns in recent data to forecast what may come next.

"This technique allows data systems to peek into the future and optimize themselves on the fly," said Aidin Niaparasat, study co-author and Ph.D. student at the Tepper School of Business at Carnegie Mellon University. "We demonstrate a clear tradeoff—the better the predictions, the faster the performance. Even when predictions are wildly off, the speed is still faster than normal."

The [software](#) is available with the supplementary material published alongside the paper; the researchers have shared their code for others to use.

The researchers say this work opens the door to further use of machine learning predictions across computer system design. Structures like search trees, hash tables, and graphs could work smarter and faster by forecasting expected data patterns. The researchers hope this inspires new ways to design algorithms and data management systems.

"Learned optimizations could lead to faster databases, improved data center efficiency, and smarter operating systems," said Benjamin Moseley, an associate professor at the Tepper School and study co-author. "We've shown predictions can beat worst-case limits. But this is just the beginning—there is enormous untapped potential in this area."

More information: Samuel McCauley et al, Online List Labeling with Predictions, *arXiv* (2023). [DOI: 10.48550/arxiv.2305.10536](https://doi.org/10.48550/arxiv.2305.10536)

Here is a [link to the poster presentation](#) for this paper.

Provided by Tepper School of Business, Carnegie Mellon University

Citation: New machine learning method predicts future data patterns to optimize data storage (2024, February 15) retrieved 27 April 2024 from <https://techxplore.com/news/2024-02-machine-method-future-patterns-optimize.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>
--