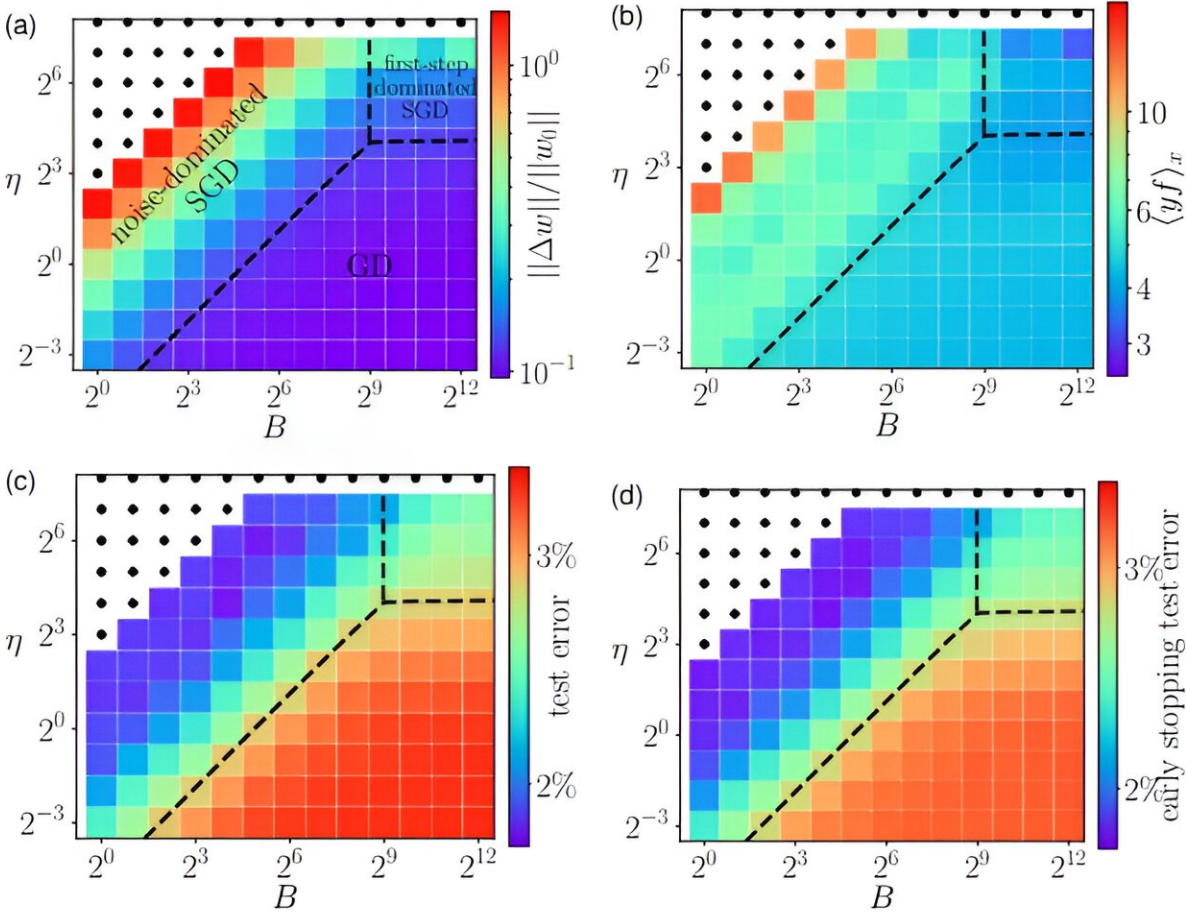


Charting new paths in AI learning: How changing two variables leads to vastly different outcomes

February 21 2024



SGD phase diagrams in batch size (B) and learning rate (η) for fully-connected (5-hidden layers) on parity MNIST. Credit: *Proceedings of the National Academy of Sciences* (2024). DOI: 10.1073/pnas.2316301121

In an era where artificial intelligence (AI) is transforming industries from health care to finance, understanding how these digital brains learn is more crucial than ever. Now, two researchers from EPFL, Antonia Sclocchi and Matthieu Wyart, have shed light on this process, focusing on a popular method known as Stochastic Gradient Descent (SGD).

At the heart of an AI's [learning process](#) are algorithms: sets of rules that guide AIs to improve based on the data they're fed. SGD is one of these algorithms, like a guiding star that helps AIs navigate a complex landscape of information to find the best possible solutions a bit at a time.

However, not all learning paths are equal. The EPFL [study](#), published in *Proceedings of the National Academy of Sciences* reveals how different approaches to SGD can significantly affect the efficiency and quality of AI learning. Specifically, the researchers examined how changing two key variables can lead to vastly different learning outcomes.

The two variables were the size of the data samples the AI learns from at a single time (this is called the "batch size") and the magnitude of its learning steps (this is the "learning rate"). They identified three distinct scenarios ("regimes"), each with unique characteristics that affect the AI's learning process differently.

In the first scenario, like exploring a new city without a map, the AI takes small, random steps, using small batches and high learning rates, which allows it to stumble upon solutions it might not have found otherwise. This approach is beneficial for exploring a wide range of possibilities but can be chaotic and unpredictable.

The second scenario involves the AI taking a significant initial step based on its first impression, using larger batches and learning rates, followed by smaller, exploratory steps. This regime can speed up the

learning process but risks missing out on better solutions that a more cautious approach might discover.

The third scenario is like using a detailed map to navigate directly to known destinations. Here, the AI uses large batches and smaller learning rates, making its learning process more predictable and less prone to random exploration. This approach is efficient but may not always lead to the most creative or optimal solutions.

The study offers a deeper understanding of the tradeoffs involved in training AI models, and highlights the importance of tailoring the learning process to the particular needs of each application. For example, medical diagnostics might benefit from a more exploratory approach where accuracy is paramount, while [voice recognition](#) might favor more direct learning paths for speed and efficiency.

More information: Antonio Sclocchi et al, On the different regimes of stochastic gradient descent, *Proceedings of the National Academy of Sciences* (2024). [DOI: 10.1073/pnas.2316301121](https://doi.org/10.1073/pnas.2316301121)

Provided by Ecole Polytechnique Federale de Lausanne

Citation: Charting new paths in AI learning: How changing two variables leads to vastly different outcomes (2024, February 21) retrieved 27 April 2024 from <https://techxplore.com/news/2024-02-paths-ai-variables-vastly-outcomes.html>

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.