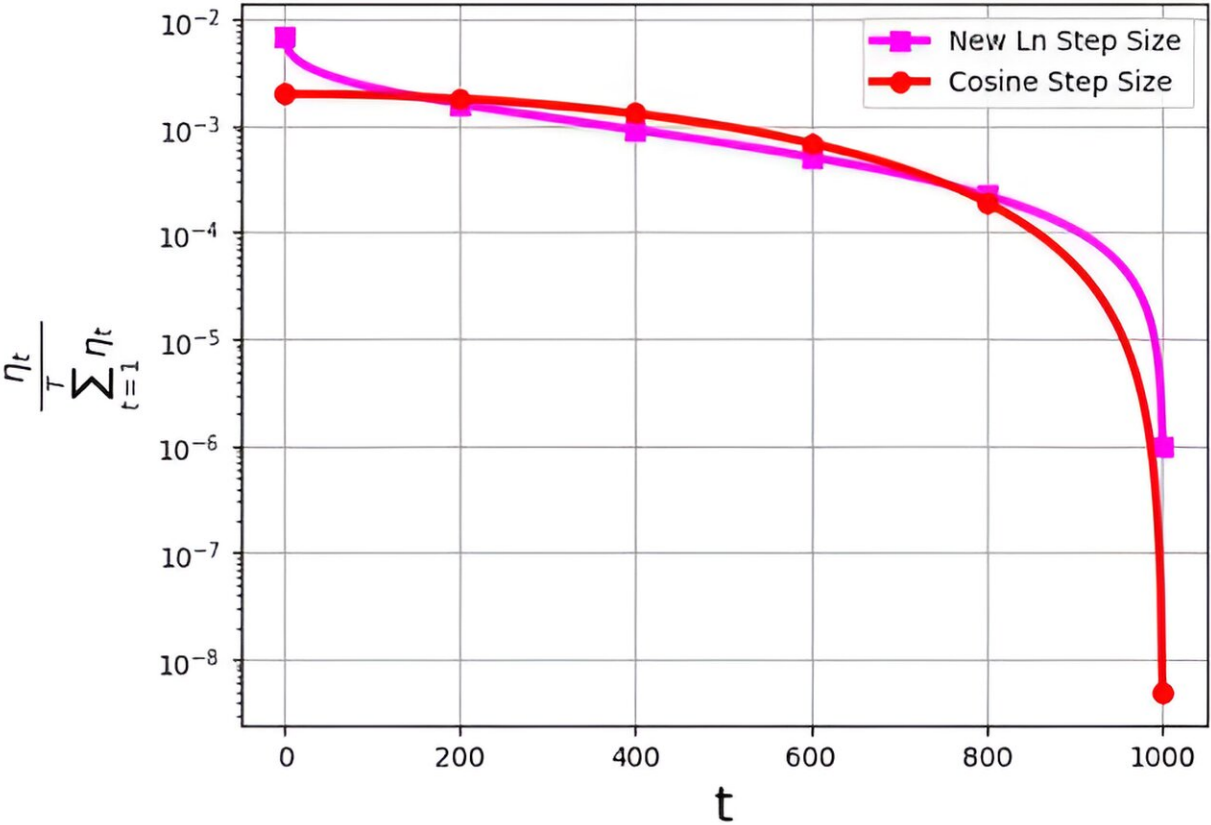


New logarithmic step size for stochastic gradient descent

April 22 2024



Credit: M. Soheil Shamaee, S. Fathi Hafshejani, Z. Saeidian

The step size, often referred to as the learning rate, plays a pivotal role in optimizing the efficiency of the stochastic gradient descent (SGD) algorithm. In recent times, multiple step size strategies have emerged for

enhancing SGD performance. However, a significant challenge associated with these step sizes is related to their probability distribution, denoted as $\eta_t / \sum_{t=1}^T \eta_t$.

This distribution has been observed to avoid assigning exceedingly small values to the final iterations. For instance, the widely used cosine step size, while effective in [practice](#), encounters this issue by assigning very low [probability distribution](#) values to the last iterations.

To address this challenge, a research team led by M. Soheil Shamaee published their [research](#) in *Frontiers of Computer Science*.

The team introduces a new logarithmic step size for the SGD approach. This new step size has proven to be particularly effective during the final iterations, where it enjoys a significantly higher probability of selection compared to the conventional cosine step size.

As a result, the new step size method surpasses the performance of the cosine step size method in these critical concluding iterations, benefiting from their increased likelihood of being chosen as the selected solution. The obtained numerical results serve as a testament to the efficiency of the newly proposed step size, particularly on the FashionMinst, CIFAR10, and CIFAR100 datasets.

Additionally, the new logarithmic step size has shown remarkable improvements in test accuracy, achieving a 0.9% increase for the CIFAR100 dataset when utilized with a convolutional neural network (CNN) model.

More information: New logarithmic step size for stochastic gradient descent, *Frontiers of Computer Science* (2024). [DOI: 10.1007/s11704-023-3245-z](https://doi.org/10.1007/s11704-023-3245-z). journal.hep.com.cn/fcs/EN/10.1007/s11704-023-3245-z

Provided by Higher Education Press

Citation: New logarithmic step size for stochastic gradient descent (2024, April 22) retrieved 20 July 2024 from <https://techxplore.com/news/2024-04-logarithmic-size-stochastic-gradient-descent.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.