A Robust Optimization Method for Label Noisy Datasets Based on Adaptive Threshold: Adaptive-*k*

Enes DEDEOGLU, Himmet Toprak KESGIN, Mehmet Fatih AMASYALI

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2430-4

Problems & Ideas

- Problems with training deep learning models in the presence of label noise:
  - Label noise in datasets can significantly decrease classification performance.
  - Existing methods for handling label noise often require prior knowledge of the dataset's noise ratio or additional model training.
- Introducing Adaptive-k, an innovative method enhancing optimization in noisy datasets without requiring prior knowledge, extra model training, or substantial computational power.

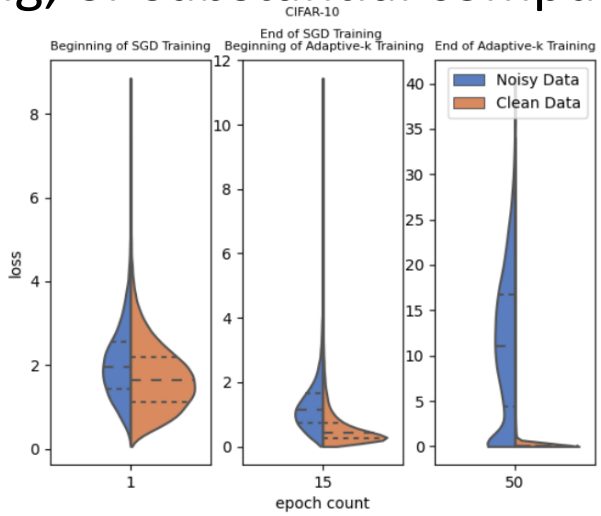
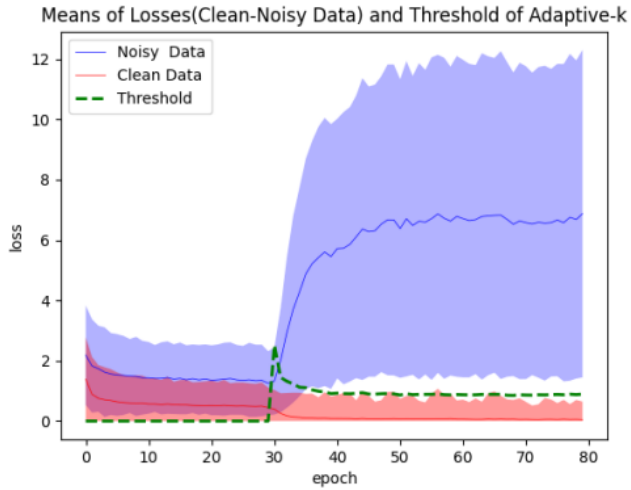


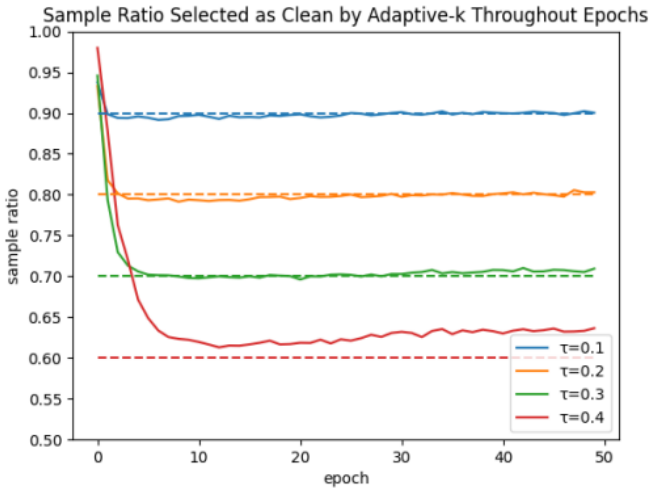
Fig. 8 Loss Distributions of Clean and Noisy Data Throughout Training for Noise Ratio $\tau = 0.3$

Main Contributions

- Contributions:
  - Adaptive-k: A novel algorithm for robust training of label noisy datasets that is easy to implement and does not require additional model training or data augmentation.
  - Theoretical analysis of Adaptive-k and comparison with the MKL algorithm and SGD.
  - High accuracy noise ratio estimation using Adaptive-k without prior knowledge of the dataset or hyperparameter adjustments.
  - Empirical comparisons of Adaptive-k with Oracle, Vanilla, MKL, Vanilla-MKL, and Trimloss algorithms on multiple image and text datasets.



Adaptive-k's effectiveness in separating clean and noisy samples using adaptive thresholds on the MNIST dataset.



demonstrates how Adaptive-k accurately estimates clean sample ratios during adaptive training for the MNIST dataset.