
[Reproducibility Report] Normalized Loss Functions for Deep Learning with Noisy Labels

Anonymous Author(s)

Affiliation

Address

email

Reproducibility Summary

1

2 *The central claim of paper is that Normalized Loss functions called "Active Passive Loss" perform better on datasets*
3 *with noisy labels. We have tried to reproduce the result on one of the metric on CIFAR-10 dataset. On training on 120*
4 *epochs and same hyper parameters mentioned on the paper, the result obtained has similar testing accuracy claimed*
5 *in the paper.*

6 Scope of Reproducibility

7 Scope of reproducibility is to train the CIFAR-10 dataset with noisy labels of 4% for the metric NCE+RCE and check
8 the claim of the test accuracy of the metric NCE+RCE mentioned in the paper.

9 1 Methodology

10 We have used author's code. It took around 23 hours to train the model for 120 epochs.

11 1.1 Datasets

12 Data used was CIFAR-10. We obtained the dataset by modifying the author's code to download it onto the local drive.

13 1.2 Hyperparameters

14 We have used the hyperparameters mentioned in their code like

- 15 • Epoch:120
- 16 • Optimizer : SGD optimizer with momentum 0.9 and weight decay 0.0001
- 17 • alpha,beta : 1.0

18 1.3 Experimental setup

19 We were trying to execute the code and reproduce the results on different hardware/software environments like differ-
20 ent environments like Local CPU, Google Colab Notebook and 2 kind of AWS GPU systems with different setup. The
21 code failed to run over certain epochs in CPU machine (RAM: 16 GB)owing to the GPU requirement of the training
22 procedure. We were not able to complete the experiment on a Google Colab because the runtime exceeded 12 hours
23 limit. We successfully run the experiment of 120 epochs on AWS Deeplearning AMI machine (CPU RAM : 32 gb, 8
24 vCPU and 2 GPU - 8gb).

25 Results

26 we reproduced the accuracy to within 1% of reported value(ours - 85.97% , paper's conclusion - 86.02%) for same
27 number of epochs , with noisy labels of 4% on same data set and code, that upholds the paper's conclusion and claim
28 that NCE+RCE performs much better than baselines

29 **What was easy**

30 It was easy to run the author's code. It was well structured and easy to run in new software/hardware setup. We were
31 able to easily execute the code on different environments like Local CPU, Google Colab Notebook and 2 kind of AWS
32 GPU systems with different setup.

33 **What was difficult**

34 We had to make small changes to the code to download the required data set. We spent some time on checking the
35 necessary compute resources required by trying different machines.

36 **Communication with original authors**

37 We did not communicate with the original authors.