SAIDL Spring Assignment 2025: Core ML Report

Vinay R Jumani

31st March 2025

1 Introduction

Robustness in machine learning is crucial when handling real-world datasets prone to noisy labels. The goal of this assignment is to explore techniques to improve robustness using normalized losses and the Active-Passive Loss (APL) framework.

2 Data Preparation

To evaluate robustness, we introduce symmetric noise into the CIFAR-10 dataset. A noise rate $\eta \in [0.2, 0.8]$ is applied by randomly flipping labels. This helps assess how models perform under different noise conditions.

3 Normalized Losses

We implemented the following loss functions:

- Cross-Entropy Loss (CE)

- Focal Loss (FL)

- Normalized Cross-Entropy Loss (NCE)

- Normalized Focal Loss (NFL)

A simple CNN architecture was trained using each of these losses across different noise levels. The performance was evaluated using accuracy metrics.

4 Active-Passive Loss Framework

APL balances robustness and performance by combining an active loss (NCE) with a passive loss (Reverse Cross-Entropy). The model was trained using APL for various noise levels, and its effectiveness was compared to the previous approaches.

5   Results and Discussion

Figures 1 and 2 illustrate the impact of different loss functions on noisy CIFAR-10 data.
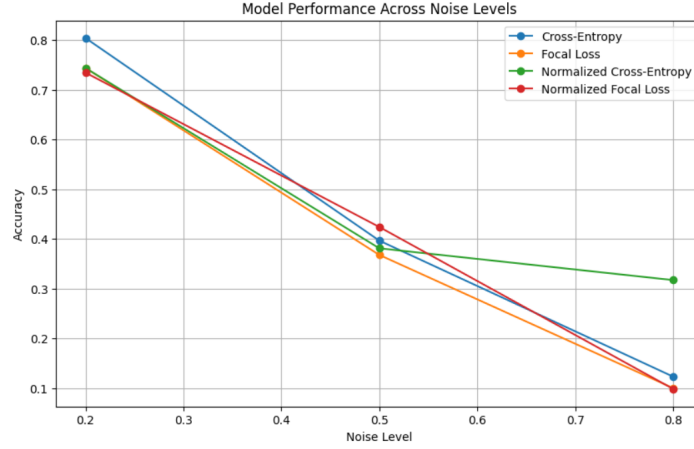

Figure 1: Performance comparison of different loss functions under varying noise levels.
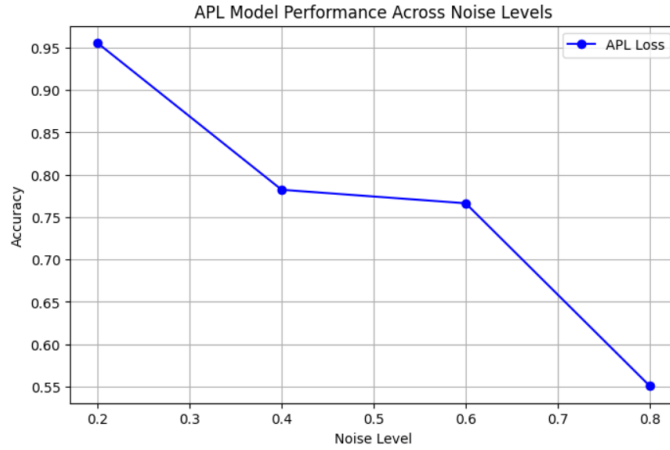

Figure 2: Performance of APL loss compared to other loss functions.

6   Conclusion

The results show that normalized losses improve robustness against label noise, and the APL framework further enhances model performance. Future work

includes testing these methods on different datasets and extending them to asymmetric noise scenarios.

7   References

References

[1] Zhang, X. et al., *Normalized Loss Functions for Deep Learning with Noisy Labels*, NeurIPS 2020.

[2] Xiao, H. et al., *Active-Passive Loss for Robust Learning*, CVPR 2023.

[3] Paszke, A. et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, NeurIPS 2019.