

Conference Paper Title*

*Note: Sub-titles are not captured in Xplore and should not be used

1 st André Lucena Ribas Ferreira	1 st Carlos Eduardo da Silva Machado	1 st Goncalo Manuel Maia de Sousa
<i>University of Minho</i>	<i>University of Minho</i>	<i>University of Minho</i>
Braga, Portugal	Braga, Portugal	Braga, Portugal
pg52672@uminho.pt	pg52675@uminho.pt	pg52682@uminho.pt

Abstract—This document is a model and instructions for **LaTeX**. This and the **IEEEtran.cls** file define the components of your paper [title, text, heads, etc.]. ***CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

Index Terms—component, formatting, style, styling, insert

I. INTRODUCTION

- Machine learning has increased in popularity
 - image classification
- studies have tried to analyse I/O patterns in DL Workflows (source)
- very few get down to kernel level
- eBPF are ...
- we seek to provide a tool to Characterize DL workloads using eBPF's

II. BACKGROUND

- DL involves iterating through a dataset
- passing it through all the layers to calculate a loss (forward pass)
- use calculated loss to update the learnable parameters of the network
- once all data is read exactly once one epoch as passed (I/O intensive)
- pytorch is a DL framework
- Distributed DNN training (data parallelism)
- checkpointing involves saving the model state
- in pytorch its done with `torch.save()` and in offitail workloads is done in between epochs
- eBPF's

III. RELATED WORK

IV. DESIGN

- Grafana

V. EVALUATION METHODOLOGY

- `dstat`, `nvidia-smi` to get cost of using the tool
- grafana dashboard to get data

VI. EVALUATION RESULTS

VII. CONCLUSION