

Michael Melesse

<https://github.com/micmelesse>

micmelesse@gmail.com

Experience

Advanced Micro Devices (AMD)
Senior Software Development Engineer

Aug 2018 - Present

- Machine Learning Kernel development
 - working on the Flash Attention Triton Backend for AMD
 - implemented the Backward Pass with features such as Variable Sequence Lengths, Multi Query Attention/ Group Query Attention, Dropout, Causal Masking, Alibi
- Machine Learning Compiler development
 - worked on AMD's backend for Triton which is OpenAI's Language/Compiler for writing Deep Learning Kernels
 - port ML operations to ROCM by generating the appropriate MLIR, LLVM IR or Assembly
- Machine Learning Framework development
 - worked on Pytorch, Tensorflow and ONNX runtime development for AMD's ROCM backend
 - developed an internal custom CI system for running Deep Learning Models on ROCM
- Machine Learning Model Performance
 - worked on identifying specific operators in models for improvement
 - worked on scaling models to run in a distributed setting
- Machine Learning Primitive development
 - worked on MIOpen which is AMD's Deep Neural Network library.
 - worked on primitives for Deep Learning models suchs RNNs, CNNs and Transformers

MedStar Health Research Institute
Software Engineer Intern

July 2016 - Aug 2016

- Applied Machine Learning models using Keras and Sci-kit to the analysis of eye-tracking data.

Education

Princeton University
Major: Computer Science B.S.E.

Sep 2014 - June 2018

Courses: Algorithms & Data Structures, Machine Learning (ML), Computer Vision, Operating Systems.

Skills

Programming Languages: Python, Triton, C++/C, CUDA/HIP, Typescript/JavaScript

Libraries: PyTorch, React, Pandas, Numpy

Tools: LLVM, MLIR, Docker, Github Actions CI, Git, Linux

Publications

- Khan, Jehandad, Paul Fultz, Artem Tamazov, Daniel Lowell, Chao Liu, **Michael Melesse**, Murali Nandhimandalam et al. "MIOpen: An Open Source Library For Deep Learning Primitives." arXiv preprint arXiv:1910.00078 (2019).