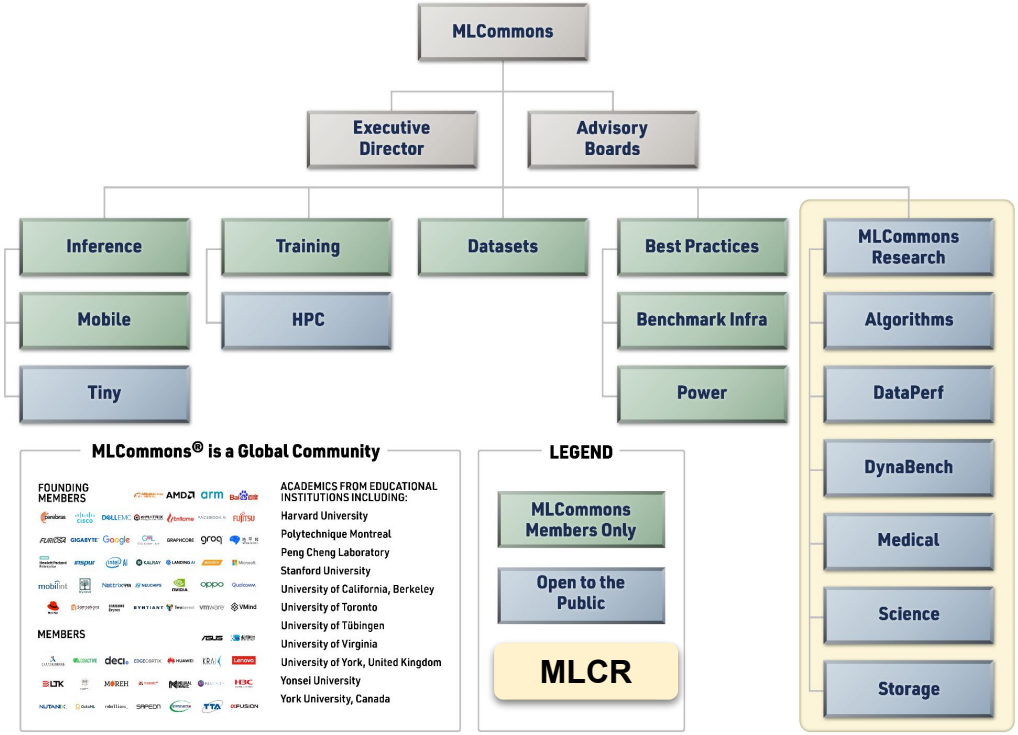

MLCommons Benchmarks
Autonomous Vehicles
A beautiful painting of ten small robots under the sun and moon in style of Monet, green color scheme

MLCommons Research Community

MLCR is **MLCommons Research** led by Janapa Reddi (Harvard) and Pekhimenko (Toronto)
Science WG is part of MLCR and is led by Fox, Hey and Thiyagalingam



Future MLCR WG's

- AR/VR (Metaverse)
- Robustness
- Autonomous Vehicles

- 62 Companies
- 11 Universities
- 6 DOE (mainly Office of Science) Labs
- ~15 WG meetings per week and
- Quarterly community meetings

- >2000 MLCommons members
- >50 FTE's from Industry
- 125 members MLCR mainly MLSys
- 27 Science WG attendees

MLPerf 2018 became **MLCommons** December 2020

"**Accelerating machine learning innovation to benefit everyone**"

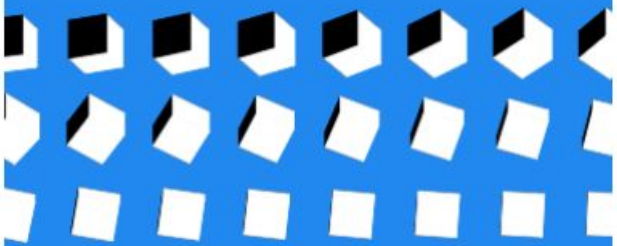
- "Grow ML markets and make the world a better place";
- "Get everyone involved";
- "Act through collaborative engineering";
- "Make fast but consensus-supported decisions," and
- "Build a community that people want to be part of."

MLCommons (MLPerf) Consortium Activity Areas

Benchmarking

Benchmarks provide consistent measurements of accuracy, speed, and efficiency. Consistent measurements enable engineers to design reliable products and services, and enable researchers to compare innovations and choose the best ideas to drive the solutions of tomorrow.

Datasets



Datasets are the raw materials for all of machine learning. Models are only as good as the data they are trained on. Academics and entrepreneurs in particular depend on public datasets to create new technologies and new companies.

Feb 7 2023: release of the Dollar Street Dataset, an open access diverse computer vision (CV) dataset. It is designed to reduce geographic and socioeconomic bias
38K high-quality images of 289 everyday objects

Best Practices

Best Practices empower researchers and engineers to more easily exchange models, reproduce experiments, and build applications that leverages machine learning. Improving best practices accelerates progress in, and grows the market for, machine learning.

Uniform Container Technology
MLCube
FAIR Metadata starting in Science

UVA Biocomplexity/CS

AI Benchr

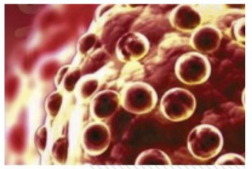
MLCommons and AI for Science

- **MLCommons** work is organized around **3 Pillars**:
  - **Best Practice** (software and ontologies)
  - **Datasets** (Identify and package good ones; develop new ones)
  - **"Benchmarks"** (the model+dataset artifacts that are open realization of mission and principles)
- **Benchmarks** measure performance OR Science & CS "Discovery"
- **Models** are critical part of infrastructure
- **DLPerf** rather than **MLPerf**: > 90% activity is **Deep learning**
- MLCommons **Best Practice WG** covers Cyberinfrastructure for managing benchmarks -- Organizing/using GitHub, logging metadata and Container infrastructure **MLCube**
- Need to add **FAIR** to MLCommons as in **FAIR for AI** meeting summary
- Provide **AI for Science** Resource in MLCommons for tutorials, demos

Science in MLCommons

- Please join and contribute to **MLCommons Research** and **Science WG**
  - https://github.com/mlcommons/science
- Improve and Add to 4 Science Benchmarks to fill out patterns of **Foundation Models**
  - Adding Surrogates from CFD and Diffusion
- Also from **MLCommons HPC** working group
  - **CosmoFlow** (3D CNN regression on cosmology simulations)
  - **DeepCAM** (2D CNN segmentation, identifying weather phenomena)
  - **OpenCatalyst** (Graph NN predicting energy and forces in atomic catalyst systems)
  - **OpenFold:** open version of protein structure prediction AlphaFold

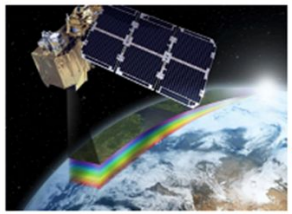
Fully Connected Network



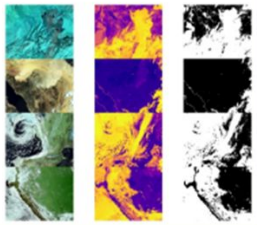
candle-uno
Deep learning-based precision medicine for cancer – drug response measurement

Uses Resnet-50 CNN

Uses U-Net CNN



slstr_cloud
Cloud Masking

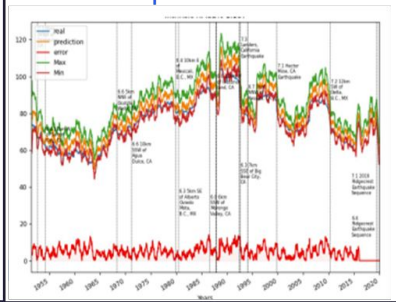


Identifying pixels that are cloud in satellite images

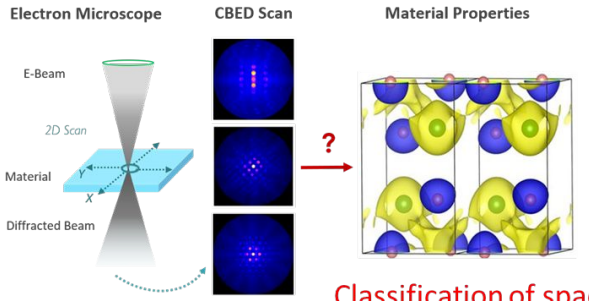
Uses Fusion Transformer + RNN

Tevelop
Timeseries Evolution Operator

Physics suggests best data processing



Focusses on extracting how a given timeseries evolves



stemdl: *ImageNet for Solid Materials*

Classification of space groups, 40GB dataset with 230 spacegroups

Deep Learning and Computing

A beautiful picture of students texting on iphones seated in a circle . The glowing sun appears as a giant computer chip with clouds in the sky. in style of Monet

Autonomous Vehicles Working Group in MLCommons

Purpose:

- Develop a benchmark for a representative automotive task for both training and inference.

Goal:

- Add a training/inference multimodal 3D object detection benchmark

Status:

- Dataset, accuracy metrics and high level model are settled.
    - Waymo Open Dataset.
    - Average Precision Heading.
    - PointPainting model.
- Working implementation of PointPainting with samples of dataset.
- Need help with compute resources for full dataset.
- Leaders Tom St. John (Tesla, Cruise), Vijay Janapa Reddi (Harvard) -- authors of Chapter 35 of AI for Science Book

Getting Started

- There are multiple GitHub repositories with code for running PointPillars. Here's a link to one of them.
  - https://github.com/zhulf0804/PointPillars
-
- We're also working on implementations of PointPainting for inclusion in MLPerf Training and Inference. If your student is interested in participating, we have meetings every other Tuesday at 2 PM Eastern time.

- Dataset https://www.argoverse.org/
- ModelNet: https://modelnet.cs.princeton.edu
- KITTI: https://www.cvlibs.net/datasets/kitti/