New Machine Learning Developments in ROOT/TMVA

Kim Albertsson EISLAB, LTU / CERN
On behalf of: Sergei Gleyzer, Mammad Hajili, Marc Huwiler, Vladimir Ilievski, Lorenzo Moneta, Saurav Shekar, Akshay Vashistha, Stefan Wunch, Omar Andres Zapata Mesa

2018-07-09

Today

- ROOT / TMVA
- News!
  - Deep learning
  - Cross validation
  - BDT Parallelisation
- A look into the future

ROOT — TMVA

ROOT aims to help high-energy physics analysis by providing building blocks for

- Data processing, analysis, visualisation, storage, parallelisation and more
- https://root.cern.ch

ROOT Machine Learning tools provided through TMVA (Toolkit for MultiVariate Analysis)

- Main ML tool for HEP applications until ~2013
- Now healthy competition with non-HEP tools
- TMVA under active development

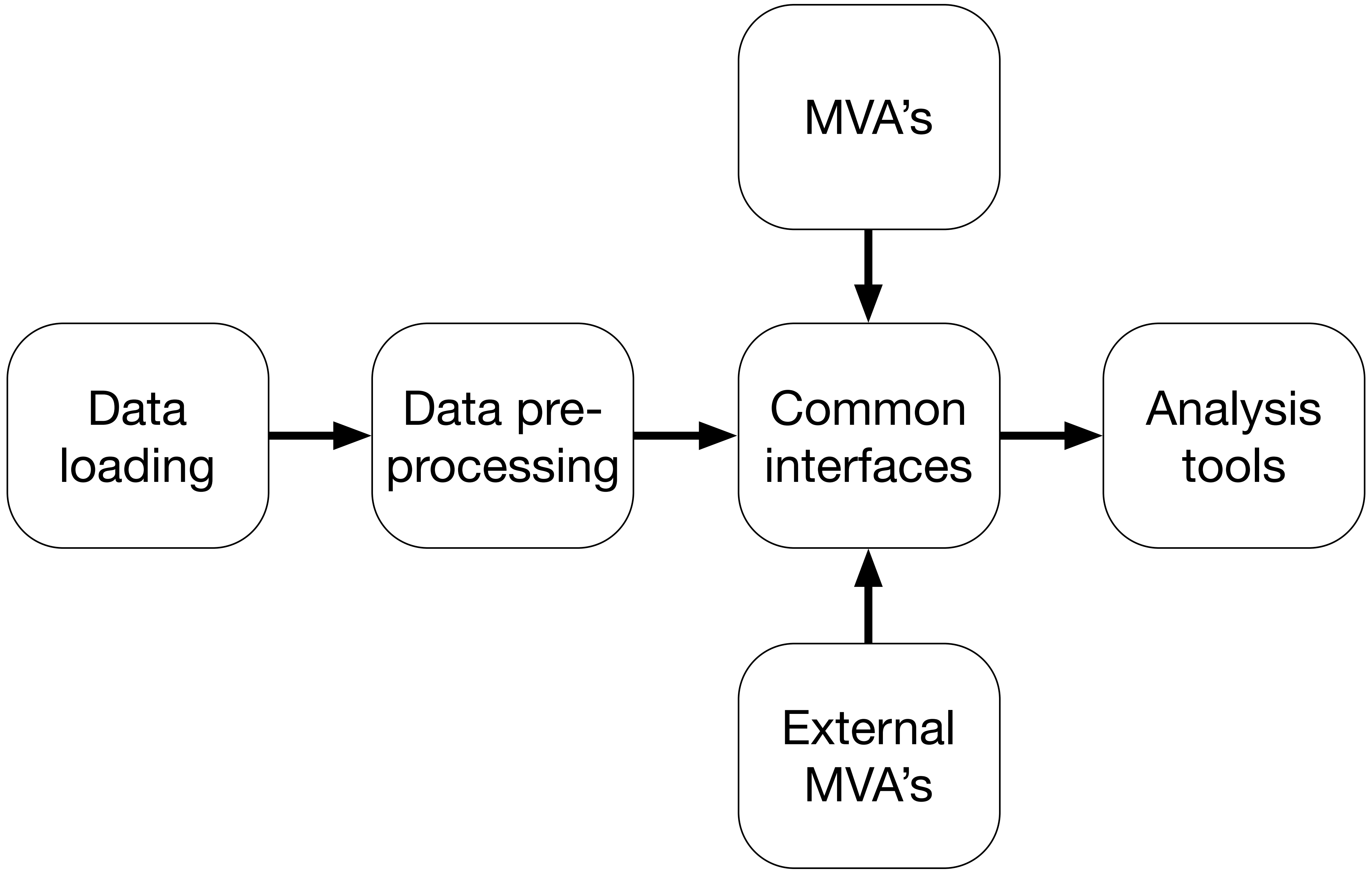
TMVA — Toolkit for multivariate analysis

TMVA is a *toolkit*

- Common interfaces
- Data loading and preprocessing
- Analysis tools

Design goals

- Easy to use
- Good out-of-the-box performance
- Standard implementation
- Targeted to HEP applications
- Long-term support



Deep Learning in HEP

Analysis

- Searching for Exotic Particles in High-Energy Physics with Deep Learning (2014) — arxiv:1402.4735
- Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the CMS detector (2018) — arxiv:1804.03682

Dense layers

Tracking/reconstruction

- Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run (2017) — ATL-PHYS-PUB-2017-013
- Jet Substructure Classification in High-Energy Physics with Deep Neural Networks (2016) — arXiv:1603.09349
- QCD-Aware Recursive Neural Networks for Jet Physics (2017) — arxiv:1702.00748
- TrackML Challenge (2018) — http://atlas.cern/updates/atlas-news/trackml-challenge

Dense layers
2D CNN
RNN

Simulation

- CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks (2017) — arXiv:1712.10321

GAN

Outside CERN

- Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber (2016) — arXiv:1611.05531

2D CNN

Plus a lot more..!

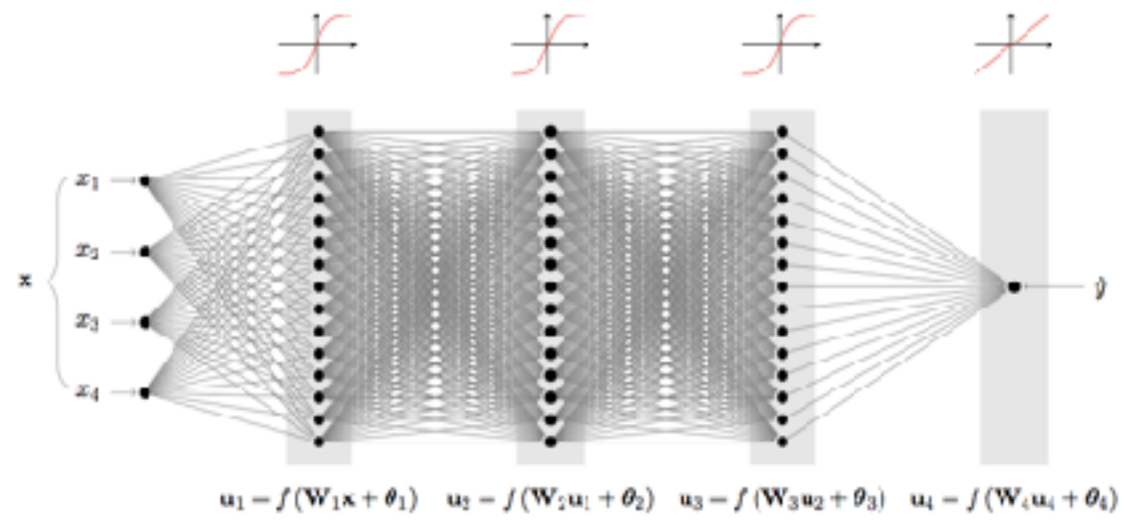
Deep Learning in TMVA

Deep learning library in TMVA

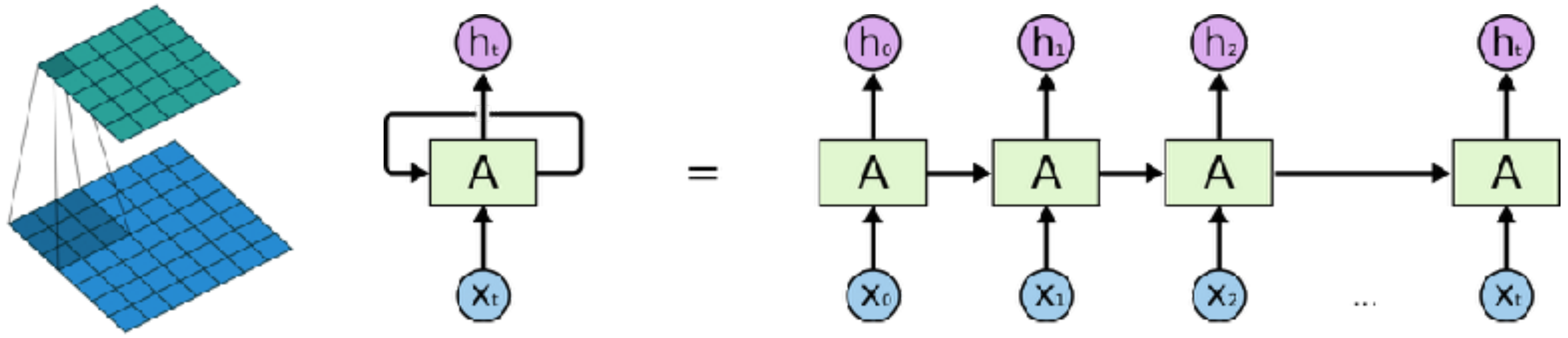
- Robust and efficient high-level DNN tools geared towards HEP
- Good out-of-the-box performance
- Do not compete with industry (proven methods!)
- CPU: BLAS + ROOT implicit multithreading (Intel TBB)
- GPU: CUDA (cuBLAS)

Timeline

- 2016 — Dense layers (CPU + GPU)
- 2017 — Convolutional, Recurrent (CPU)
- 2018 — Convolutional (GPU), Generative, LSTM (CPU)



2016

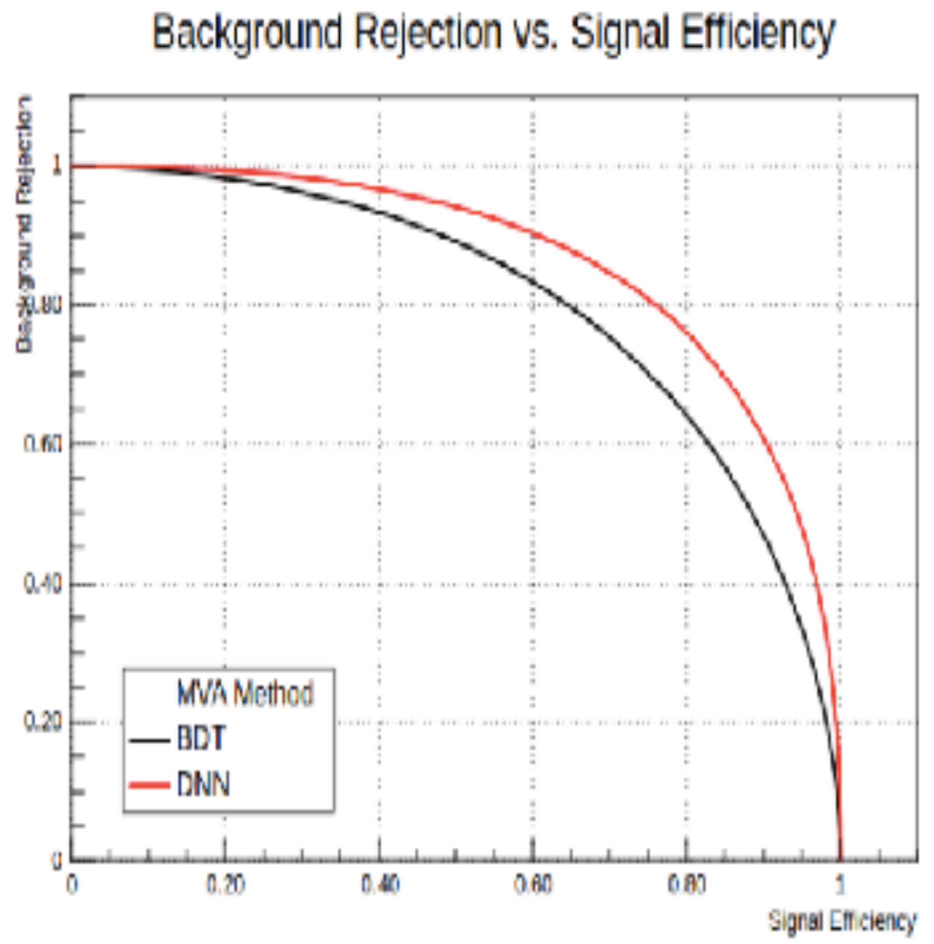


2017

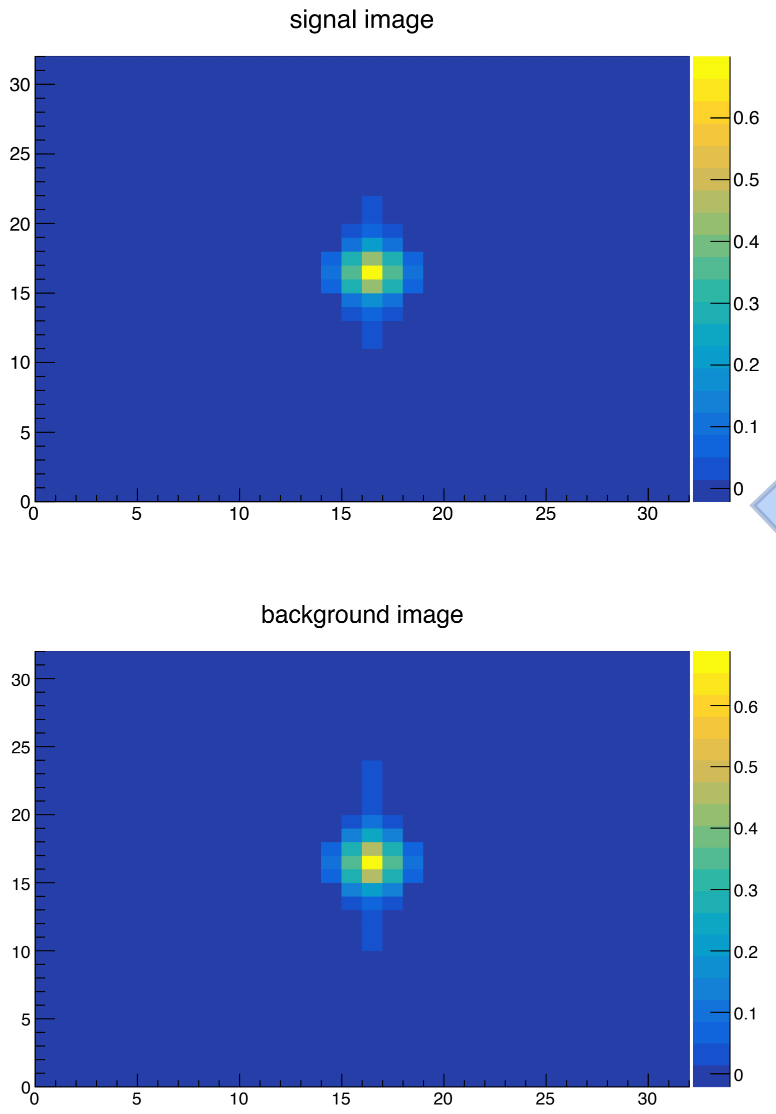
Deep Learning Performance

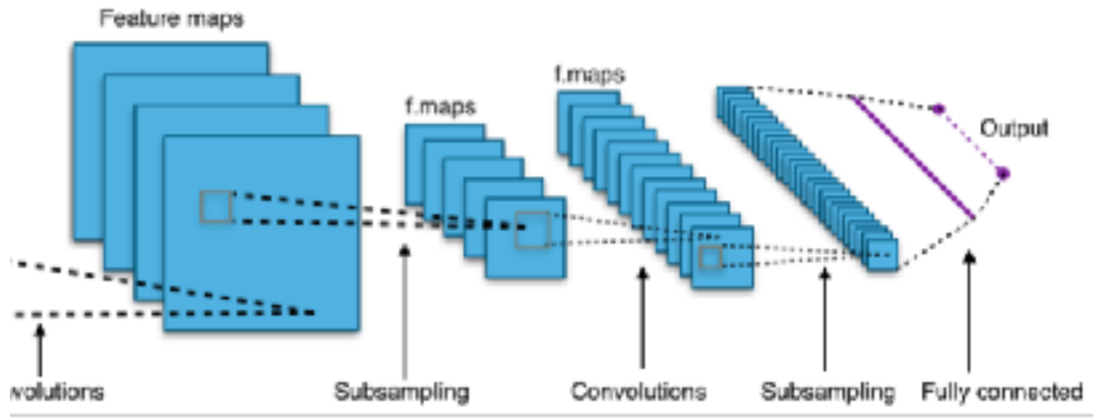
- Electron imaging in CMS calorimeter (internal)
- Identical setups with TMVA and Keras
- CNN outperforms Dense network
- Identical performance for keras and TMVA

- Using Higgs public dataset with 11M events
- Significant improvements compared to shallow networks and BDT

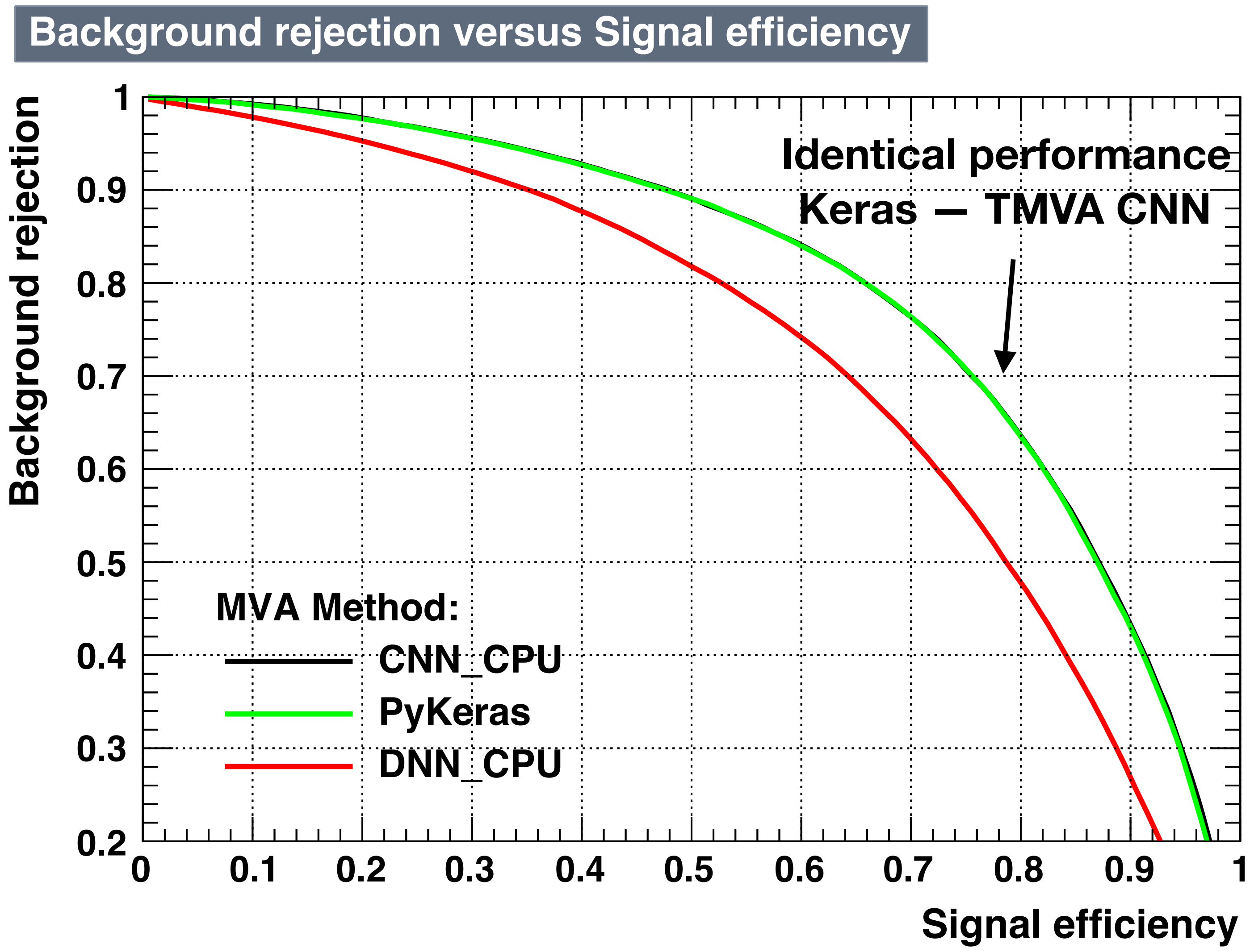


input
32x32 images





Convolutional + Pooling + Dense layers



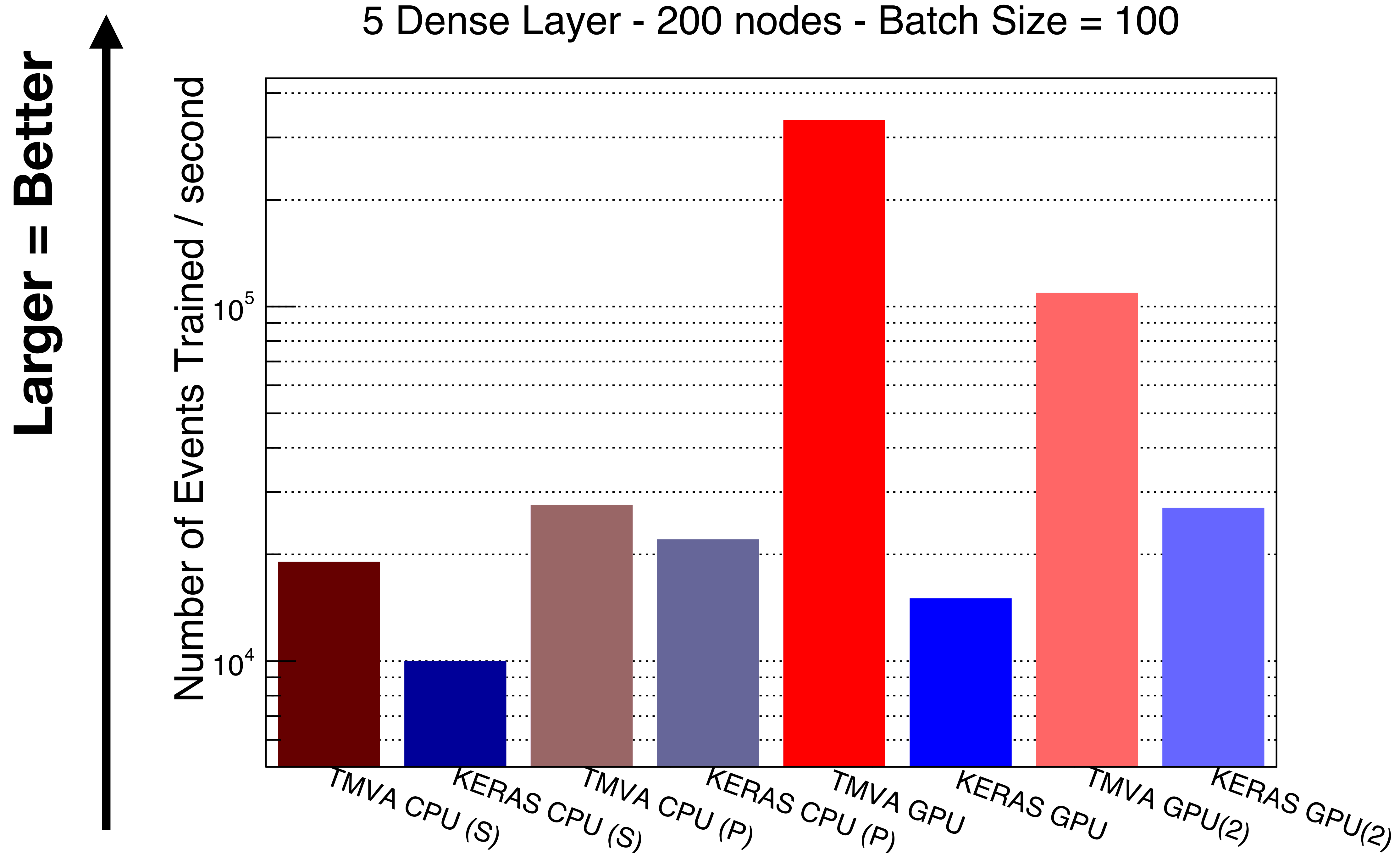
Deep Learning Performance

Training time — Dense networks

- HiggsML dataset with 11M Events
- 5 layers 200 hidden units each
- Keras with TensorFlow backend
- "Out-of-the-box" performance of TMVA vs. keras

Excellent TMVA performance!

- Better than keras?



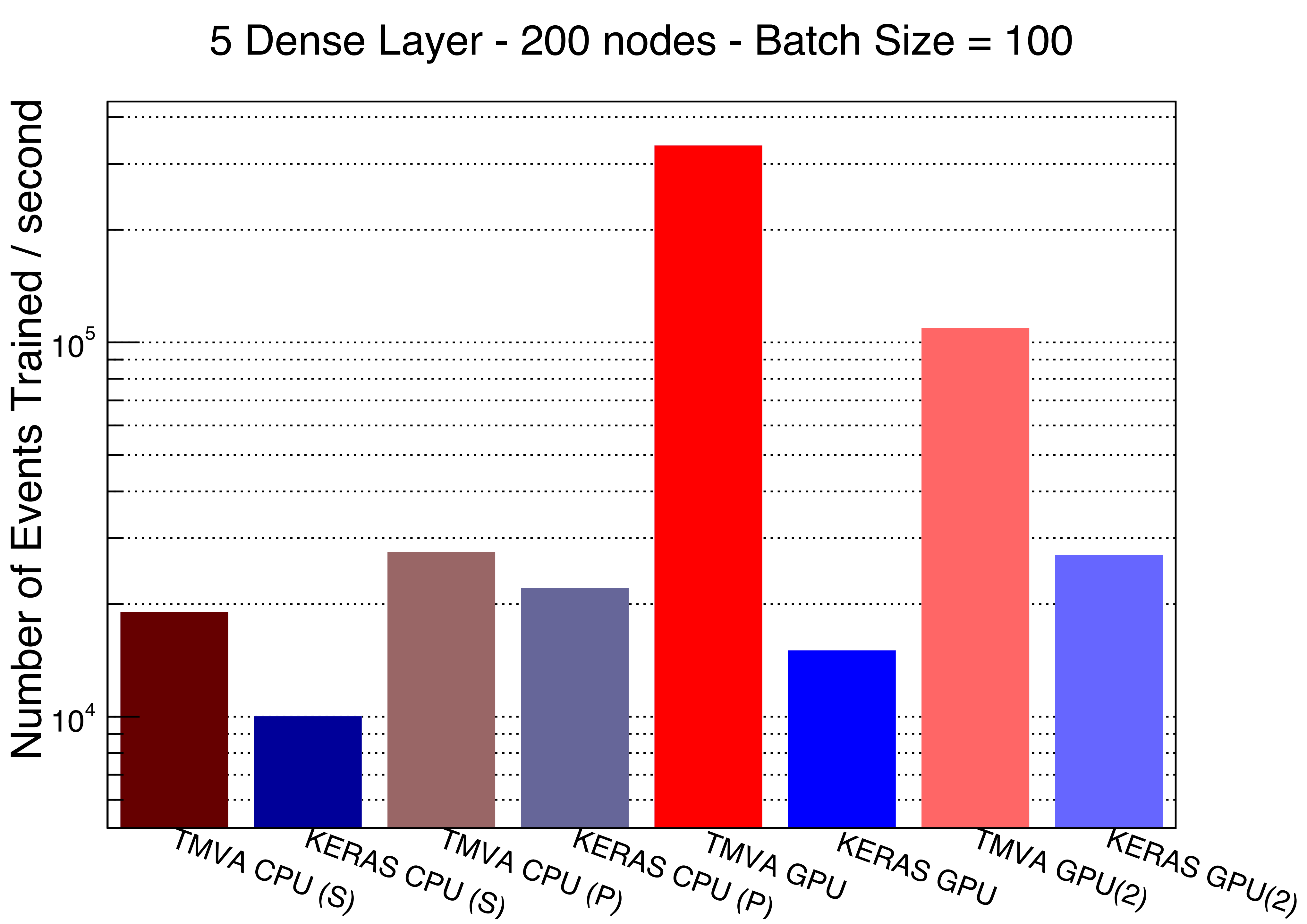
(S) — Single threaded

(P) — 32-core machine

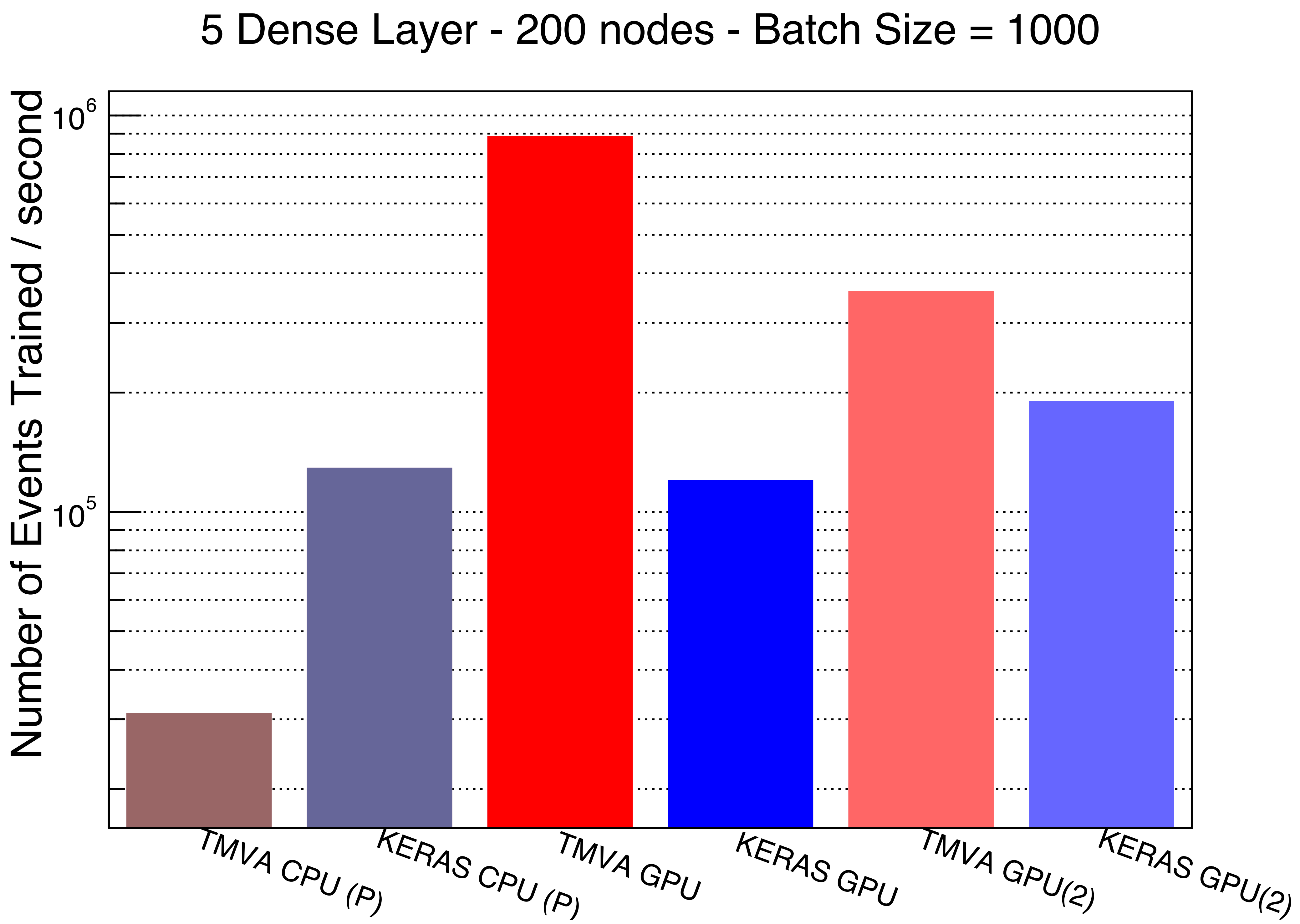
GPU — "Good" graphics card

GPU(2) — "Worse" graphics card

Deep Learning Performance



Batch size 100



Batch size 1000

NOT saying TMVA outperforms TensorFlow!

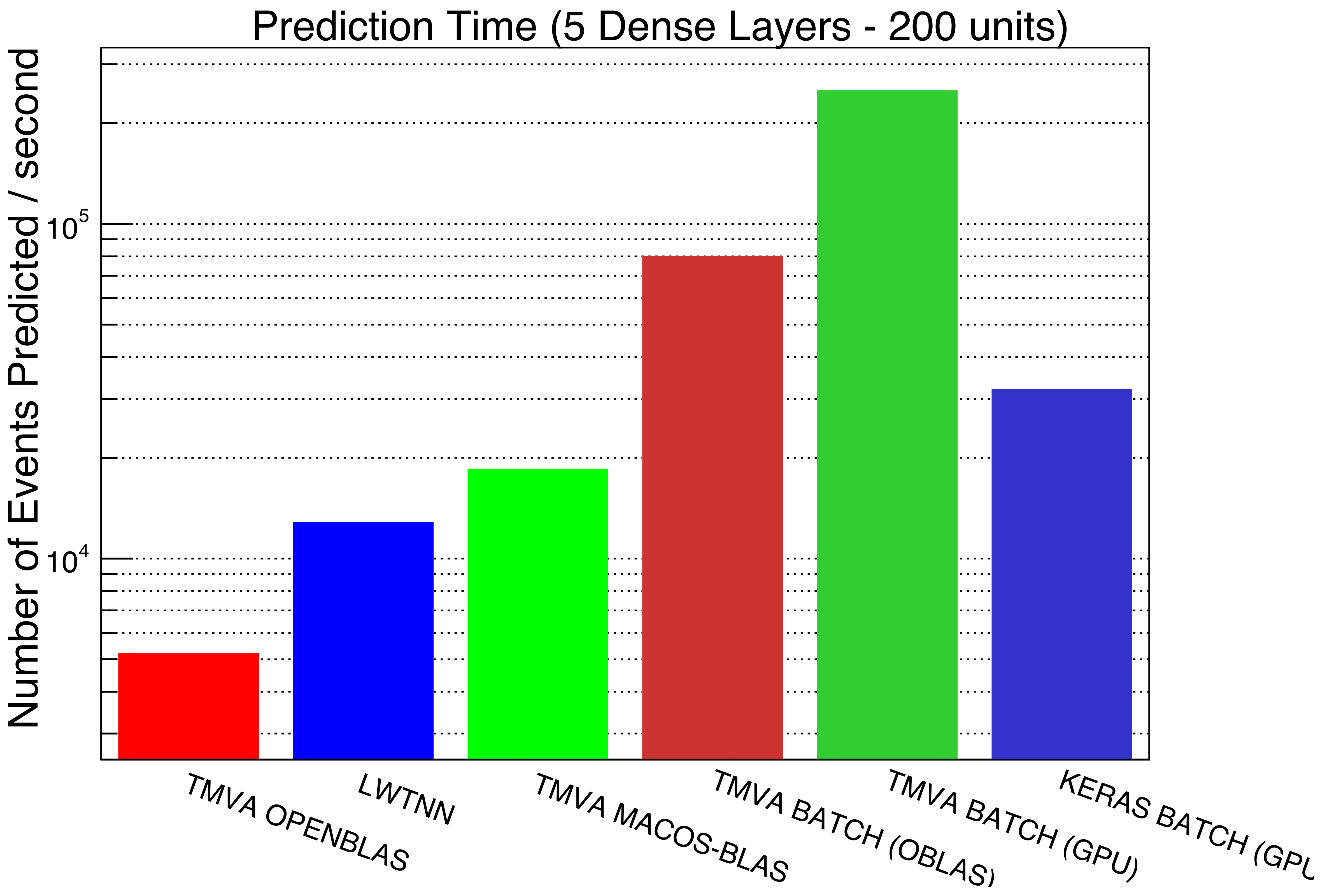
- cuBLAS for both TMVA + TF
- TF optimised for large operations

Deep Learning Performance

Evaluation time — Dense networks

- Comparison TMVA, keras, LWTNN
- In time critical applications — e.g. trigger
- Batching not an option?

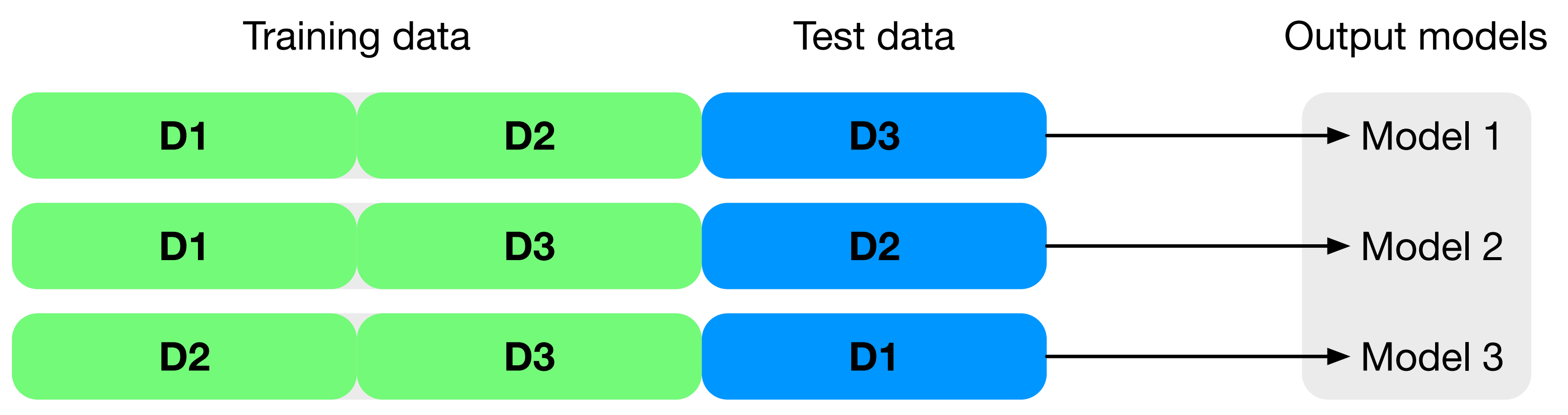
Again — keras/tensorflow benefits from larger networks



Cross Validation

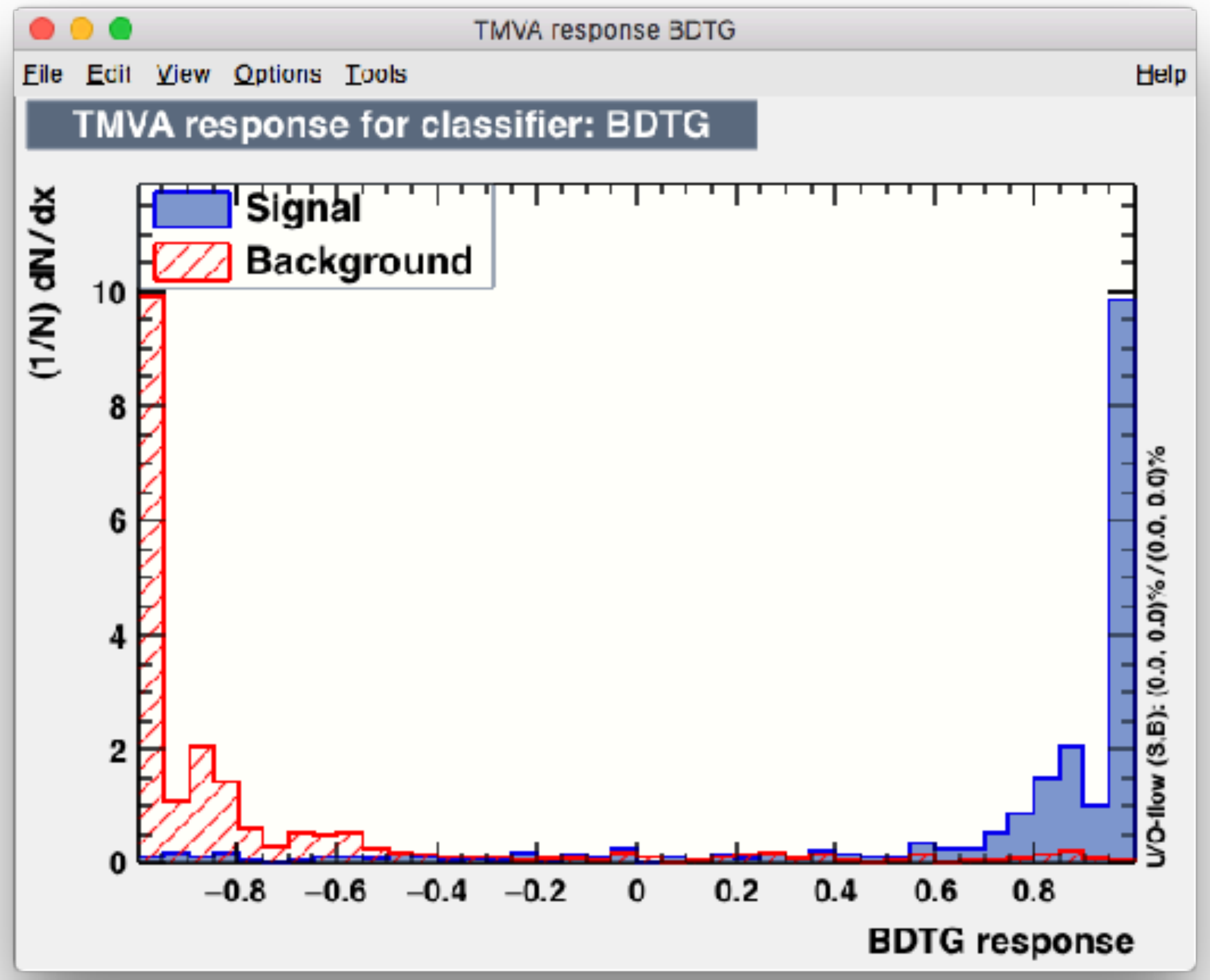
Cross validation — Efficient use of data

- Data generation generally expensive in HEP
- Deep models are large; risk of overtraining
- TMVA implements K-folds



Integration with TMVA workflow

- Now possible to use with TMVA analysis tools
- Natively supports "CV in application"



Parallel evaluation of folds through ROOT multi processing!

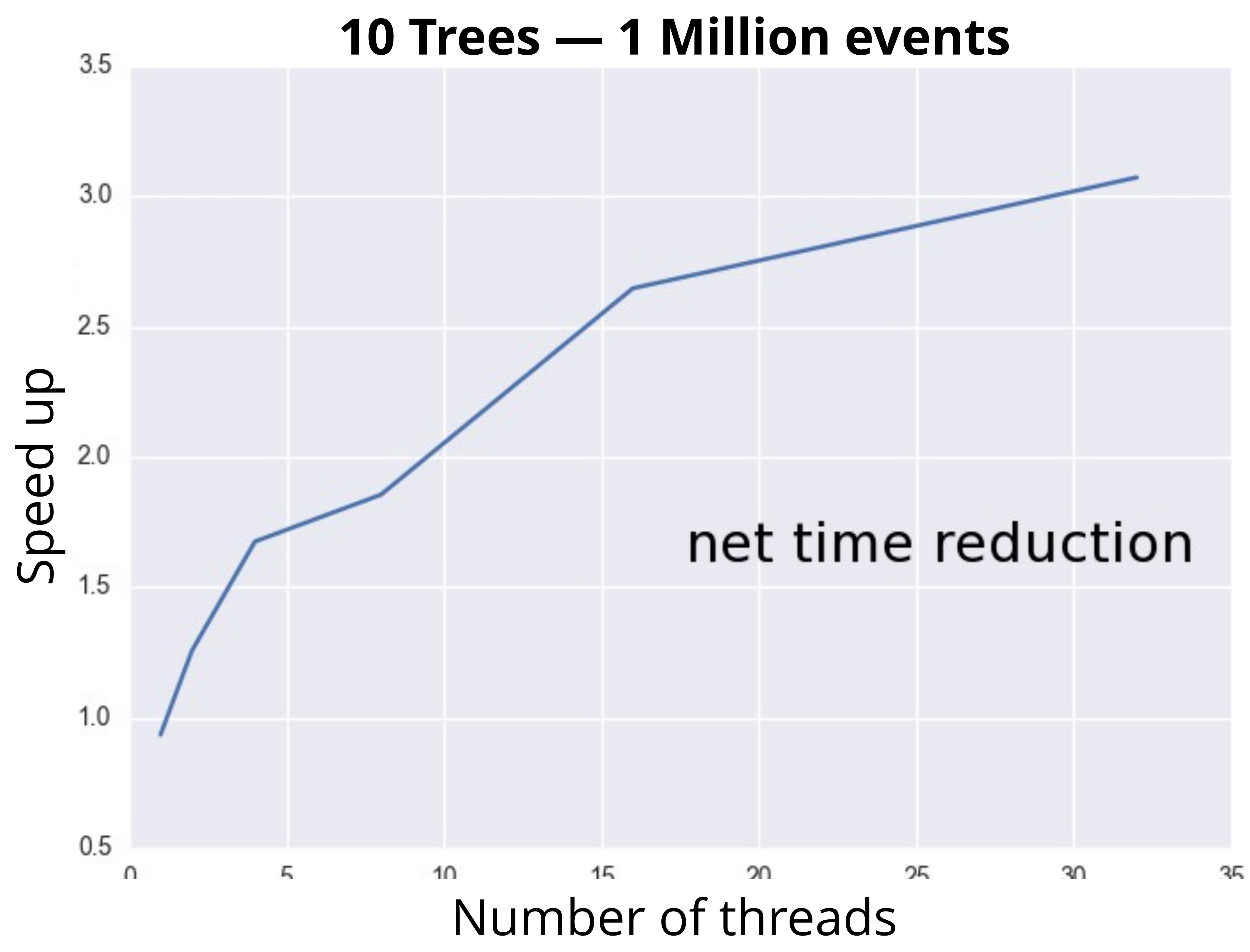
Boosted Decision Tree

Good performance over a wide variety of problems

- Used extensively in Run 1 and 2 in trigger, reconstruction and analysis
- Popular outside of HEP world

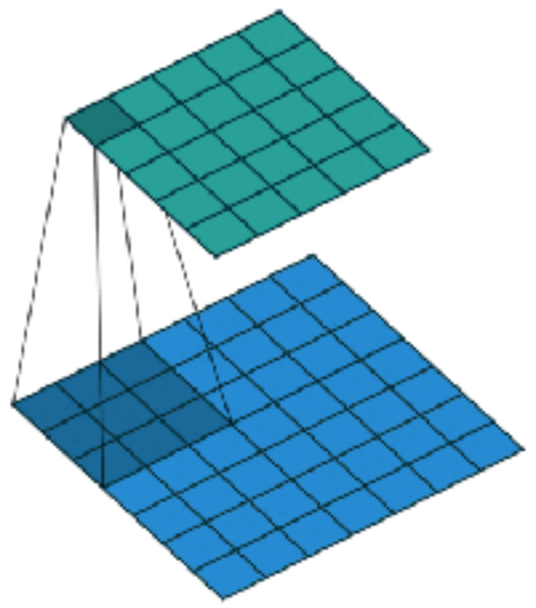
Efforts to parallelise BDT's

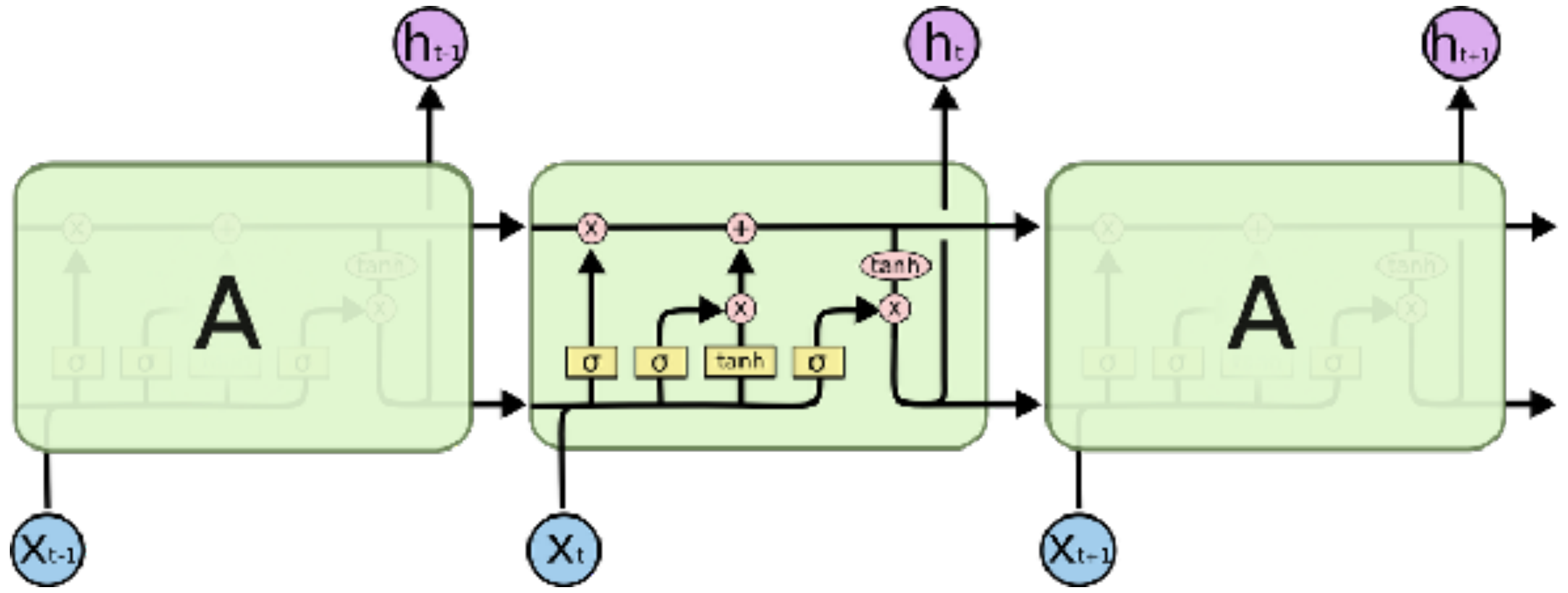
- Speed up ~1.4 with 4 threads approaching ~3 asymptotically

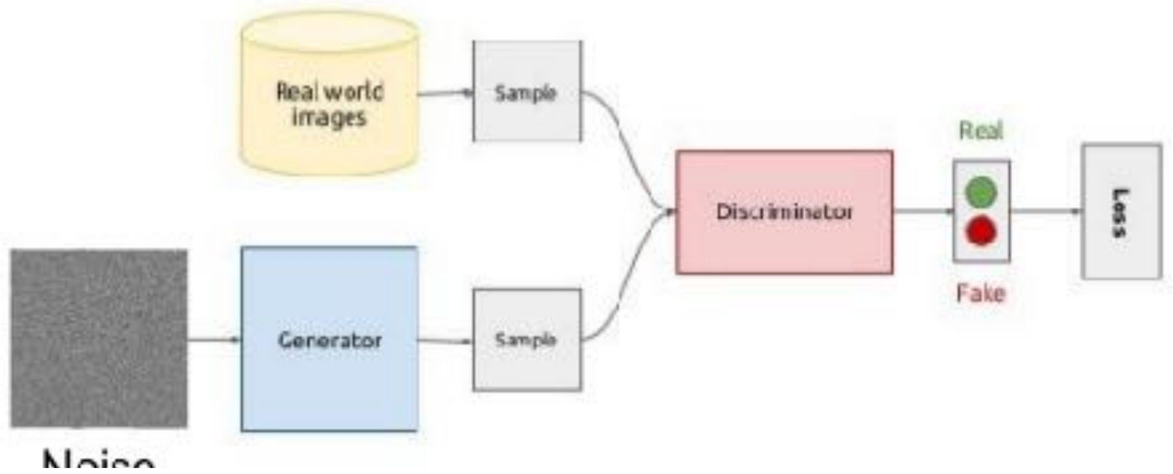


Future — Short term

- DNN Library
  - GPU: CNN
  - CPU: LSTM, GAN, VAE
  - Optimisers
- CV for model selection
- Modernised interfaces, integration with ROOT RDataFrame

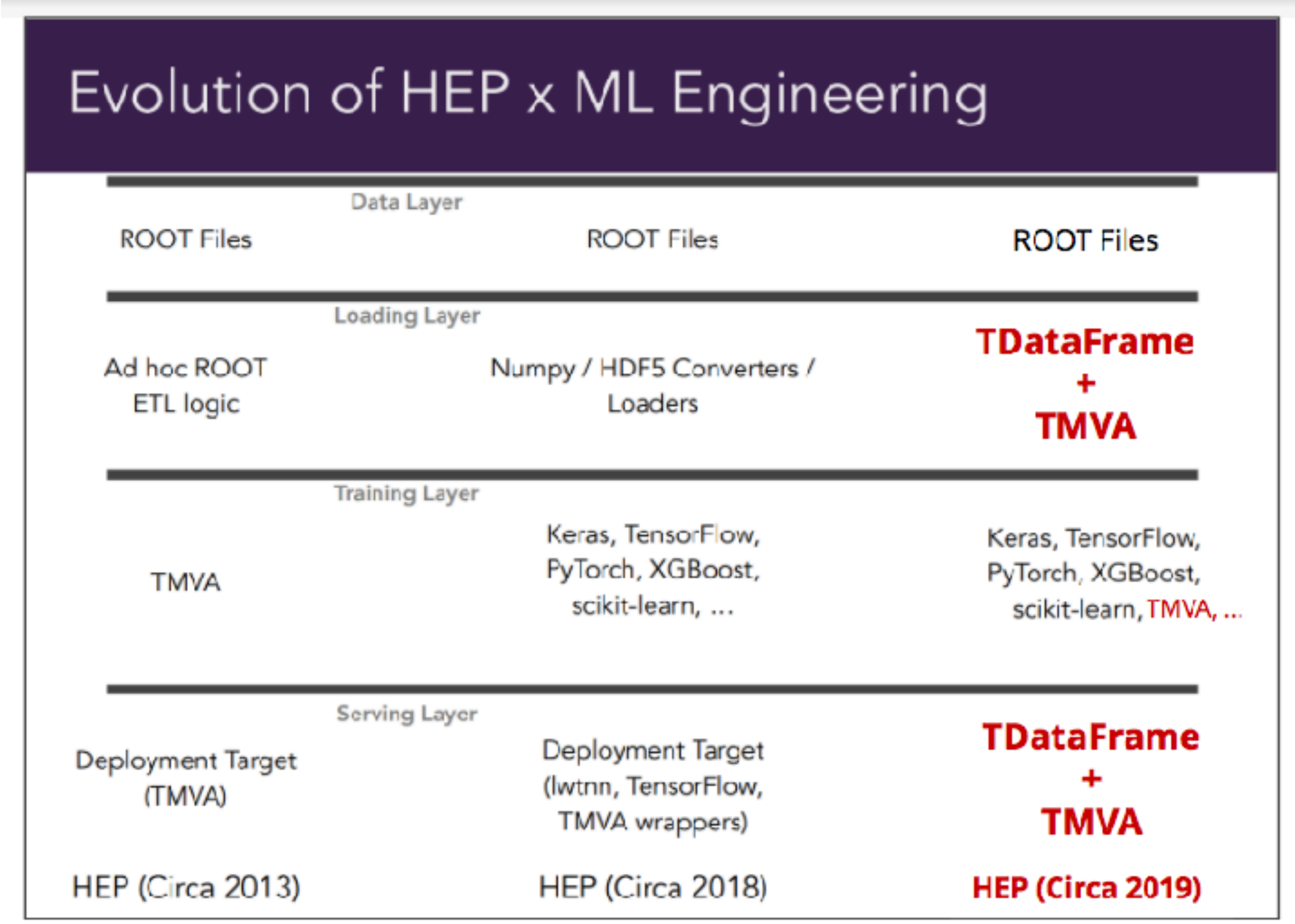






Future — Long term

- Lesson from the HSF Community white paper — Efficient workflows
- Focus on toolkit part, provide tools for efficient
  - data loading (for training)
  - integration with external tools
  - deployment



Original slide by Luke de Oliveira — Adapted by Stefan Wunsch

TDataFrame now called RDataFrame (https://root.cern/blog/rootrwhy)

Thanks

Get in touch

https://root.cern.ch

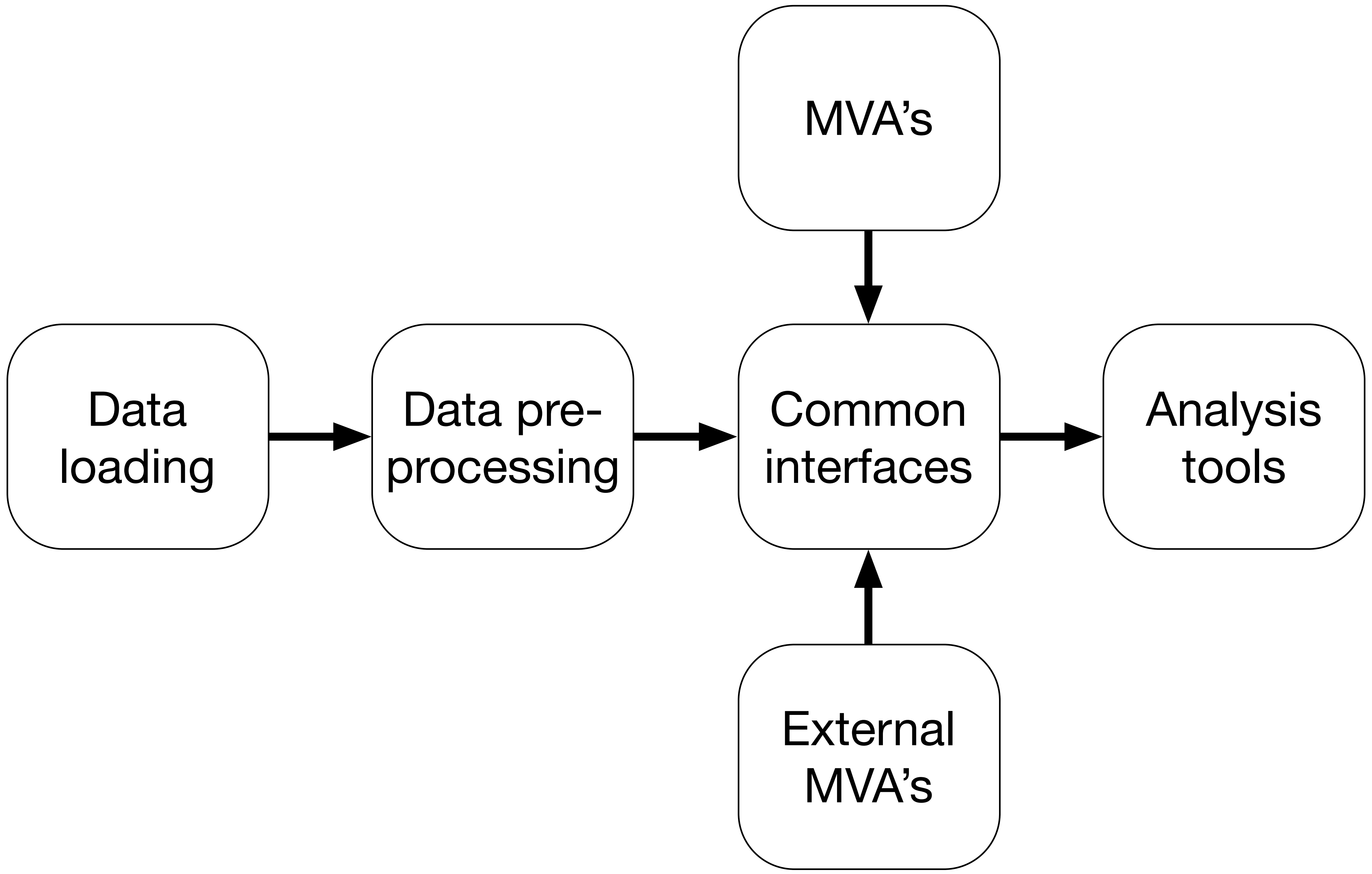
https://root-forum.cern.ch

Extra slides

TMVA — Toolkit for multivariate analysis

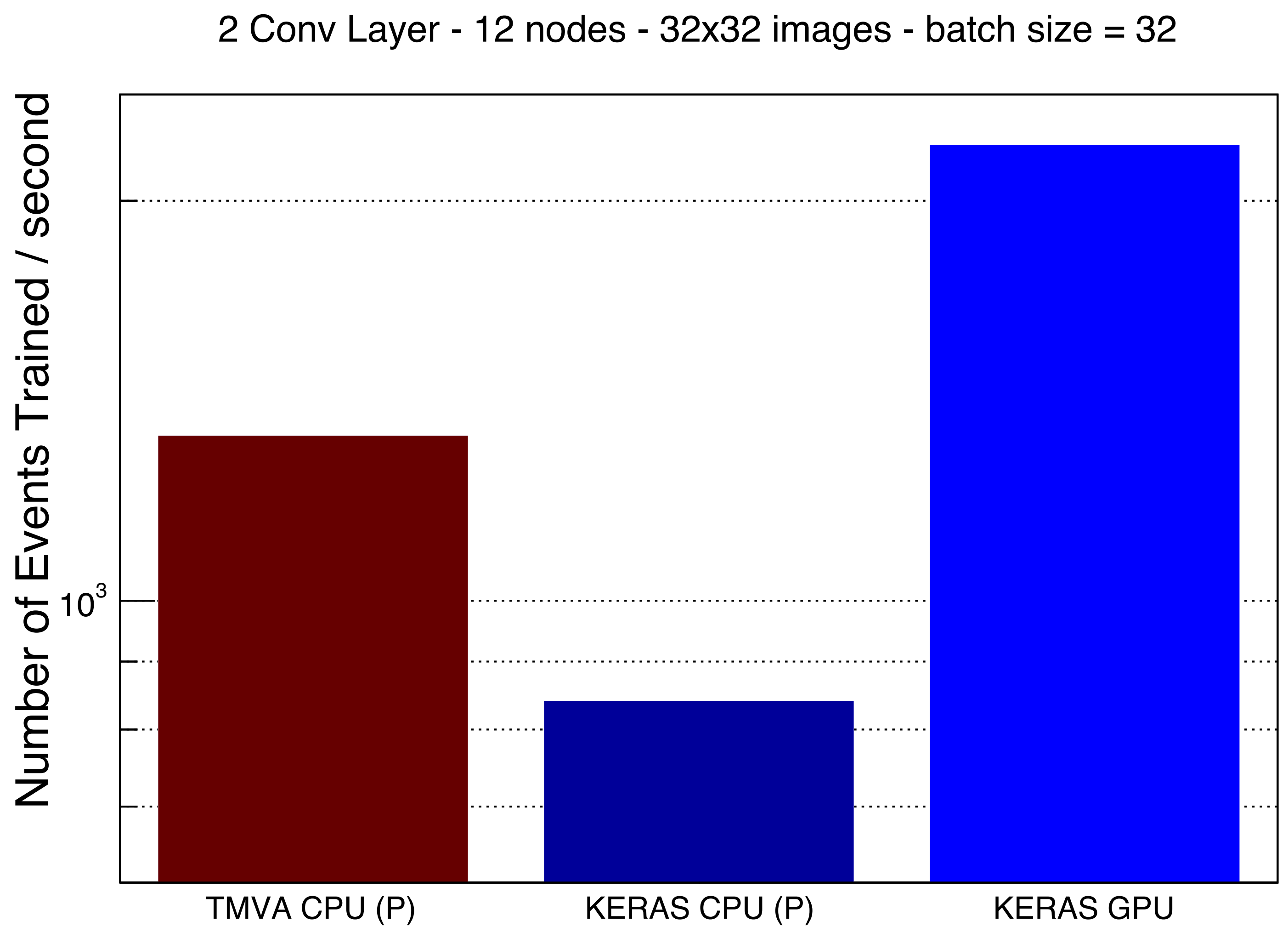
TMVA is a *toolkit*

- Common interfaces
- Data loading and preprocessing
- Analysis tools



Design goals

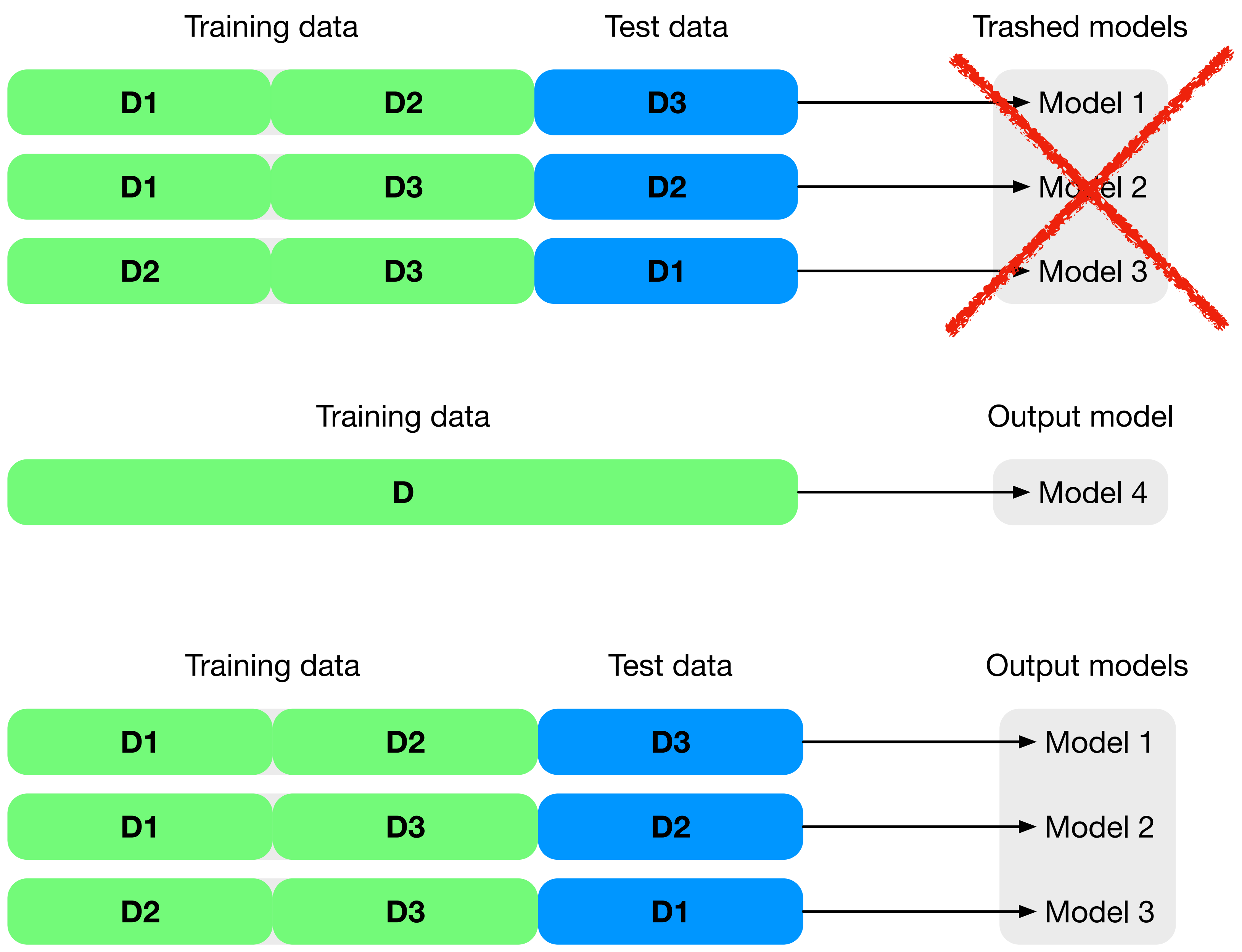
- Easy to use
- Good out-of-the-box performance
- Standard implementation
- Targeted to HEP applications
- Long-term support


2 Conv Layer - 12 nodes - 32x32 images - batch size = 32
Number of Events Trained / second
10^3
TMVA CPU (P)
KERAS CPU (P)
KERAS GPU

Cross Validation

TMVA supports “CV in application”

- Common workflow in HEP analysis
- Deterministic assignment of events to folds + save all trained models
- Performance estimation holds for collection of models



Used in e.g.

- Evidence for the H → bb̄ decay with the ATLAS detector (2017) — arxiv:1708.03299
- Search for the bb decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector (2015) — arxiv:1409.6212