Deep Learning at UHH

Gregor Kasieczka, Peter Schleper
(gregor.kasieczka@uni-hamburg.de)

ErUM-Data Kick-Off Meeting

2019-02-25

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft
DFG

Bundesministerium für Bildung und Forschung

Optimal Architectures for LHC (jet) data

Public next week:

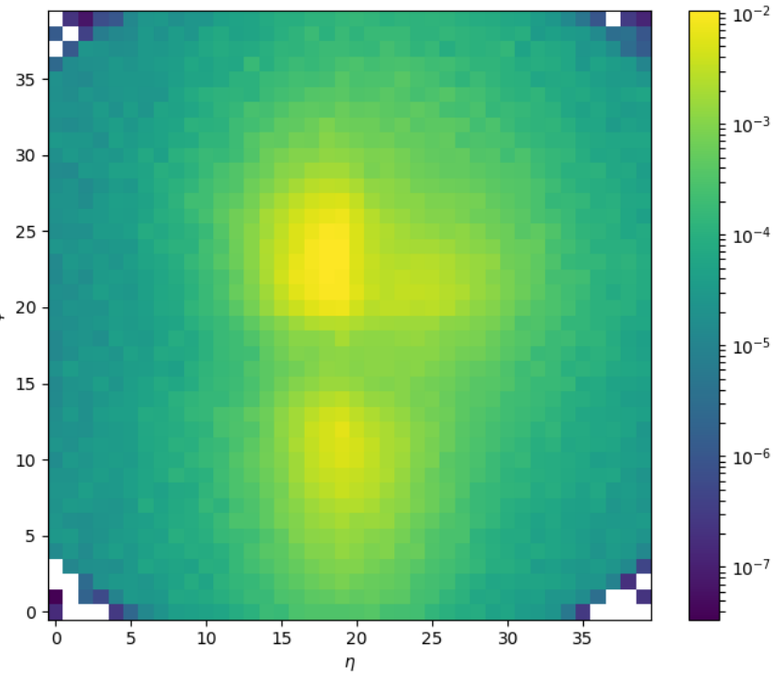
The Machine Learning Landscape of Top Taggers

G. Kasieczka (ed)[1], T. Plehn (ed)[2], A. Butter[2], D. Debnath[3], M. Fairbairn[4], W. Fedorko[5], C. Gay[5], L. Gouskos[6], P. T. Komiske[7], S. Leiss[1], A. Lister[5], S. Macaluso[3], E. M. Metodiev[7], L. Moore[8], B. Nachman,[9,10], K. Nordström[11,12], J. Pearkes[5], H. Qu[6], Y. Rath[13], M. Riegler[13], D. Shih[3], J. M. Thompson[2], and S. Varma[4]

1 Institut für Experimentalphysik, Universität Hamburg, Germany
2 Institut für Theoretische Physik, Universität Heidelberg, Germany
3 NHECT, Dept. of Physics and Astronomy, Rutgers, The State University of NJ, USA
4 Theoretical Particle Physics and Cosmology, King's College London U.K.
5 Department of Physics and Astronomy, The University of British Columbia, Canada
6 Department of Physics, University of California, Santa Barbara, USA
7 Center for Theoretical Physics, MIT, Cambridge, USA
8 CP3, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
9 Physics Division, Lawrence Berkeley National Laboratory, Berkeley, USA
10 Simons Inst. for the Theory of Computing, University of California, Berkeley, USA
11 National Institute for Subatomic Physics (NIKHEF), Amsterdam, Netherlands
12 LPTHE, CNRS & Sorbonne Université, Paris, France
13 III. Physics Institute A, RWTH Aachen University, Germany

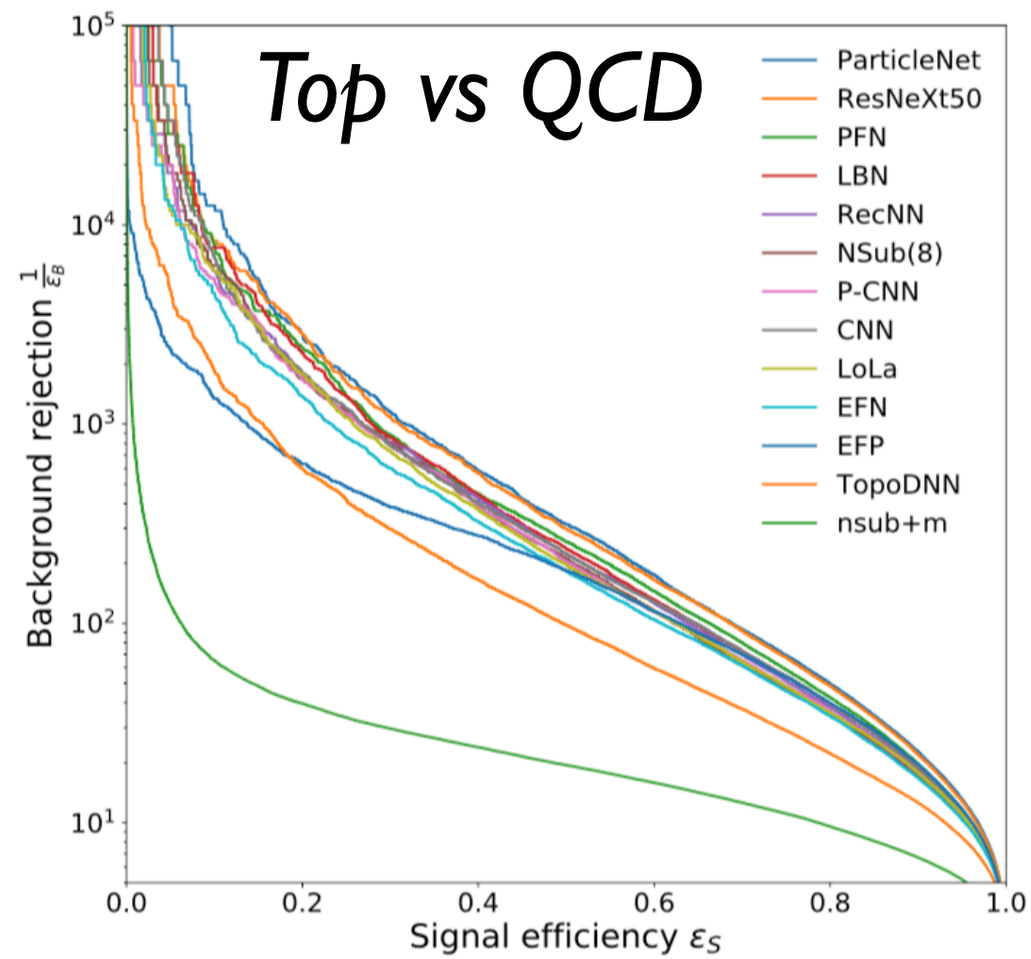
gregor.kasieczka@uni-hamburg.de

February 19, 2019



Jet Images

Deep-learning Top Taggers or The End of QCD?
GK, T Plehn, M Russell, T Schell
JHEP 05 (2017) 006



$$\tilde{k}_j \xrightarrow{\text{LoLa}} \hat{k}_j = \begin{pmatrix} m^2(\tilde{k}_j) \\ p_T(\tilde{k}_j) \\ w_{jm}^{(E)}\, E(\tilde{k}_m) \\ w_{jm}^{(p_T)}\, p_T(\tilde{k}_m) \\ w_{jm}^{(m^2)}\, m^2(\tilde{k}_m) \\ w_{jm}^{(d)}\, d_{jm}^2 \end{pmatrix}$$

4 Vectors

Deep-learning Top Taggers & No End to QCD, A Butter, GK, T Plehn, M Russell, 1707.08966, SciPost Phys. 5, 028 (2018)

$$k_{\mu,i} \xrightarrow{\text{CoLa}} \tilde{k}_{\mu,j} = k_{\mu,i}\, C_{ij} \quad \text{with} \quad C = \begin{pmatrix} 1 & \cdots & 0 & C_{1,N+1} & \cdots & C_{1,M} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & C_{N,N+1} & \cdots & C_{N,M} \end{pmatrix}$$

Work with Heidelberg (Plehn)

Tagging is essentially solved.

What's next?

Application in analysis
(Long Lived Particles, Higgs, Top)

Assigning & Reducing Uncertainties

Getting Closer to the Detector

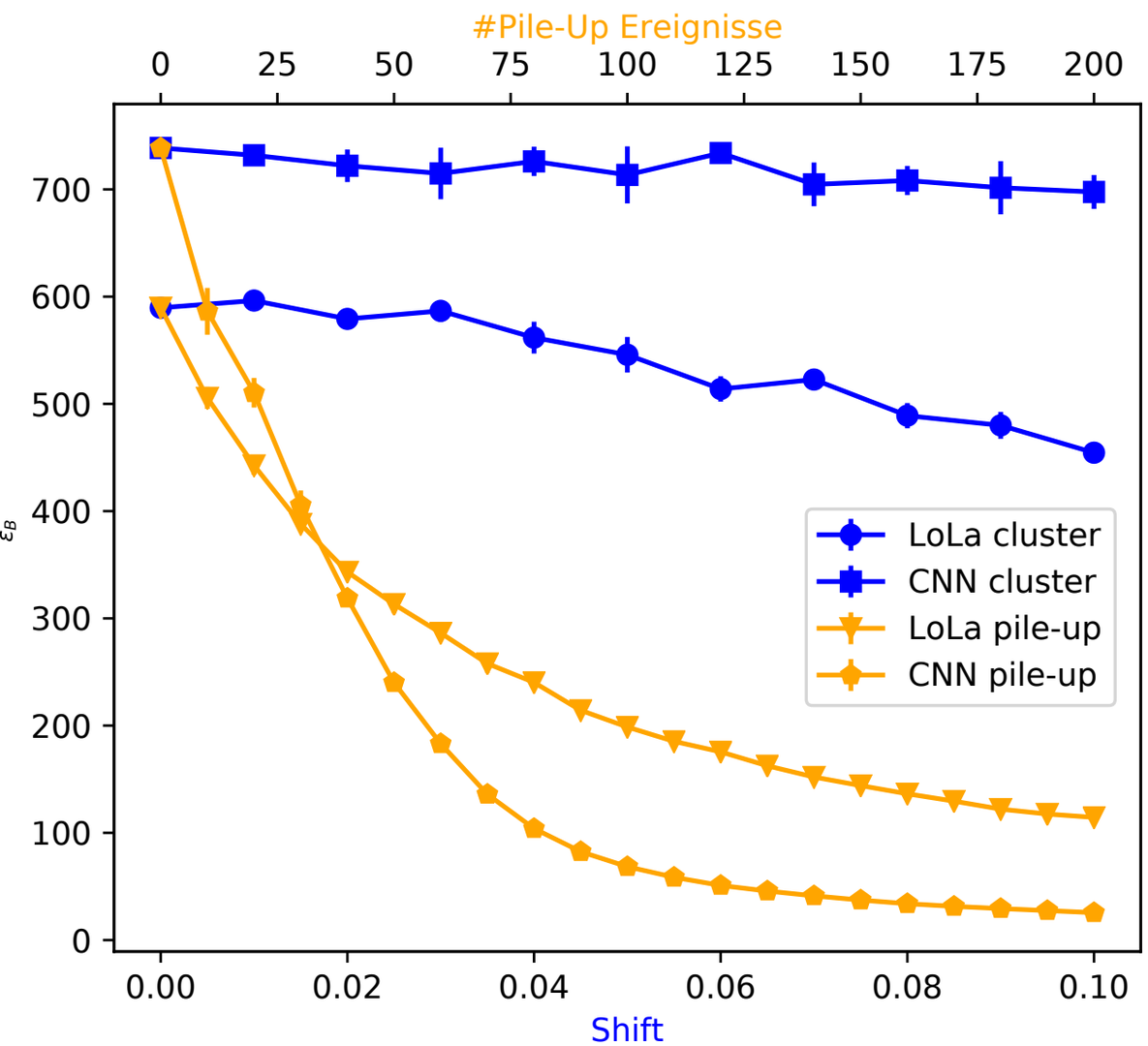
Asking new Questions

Understanding

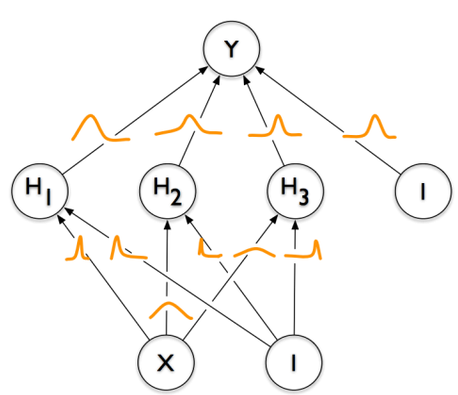
Finding new Connections

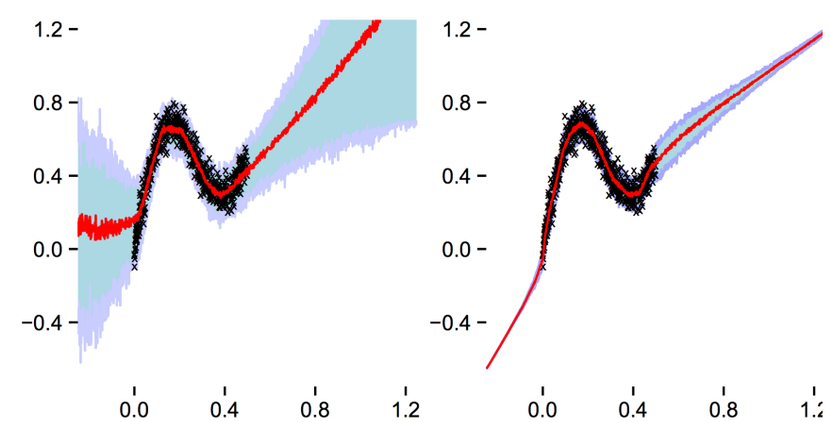
Of course there are other exciting topics as well.
This is just what UHH is working on…

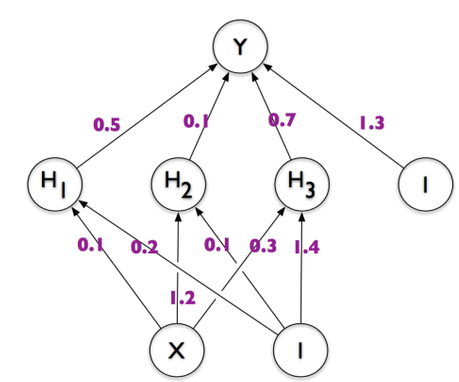
Uncertainties



- Simulate systematic differences between training MC and collision data
- Test network response under
  - rescaling of 4-vector inputs (mimic jet energy scale)
  - adding Pile-Up
- Test mitigation with data augmentation/ adversarial training/Bayesian networks







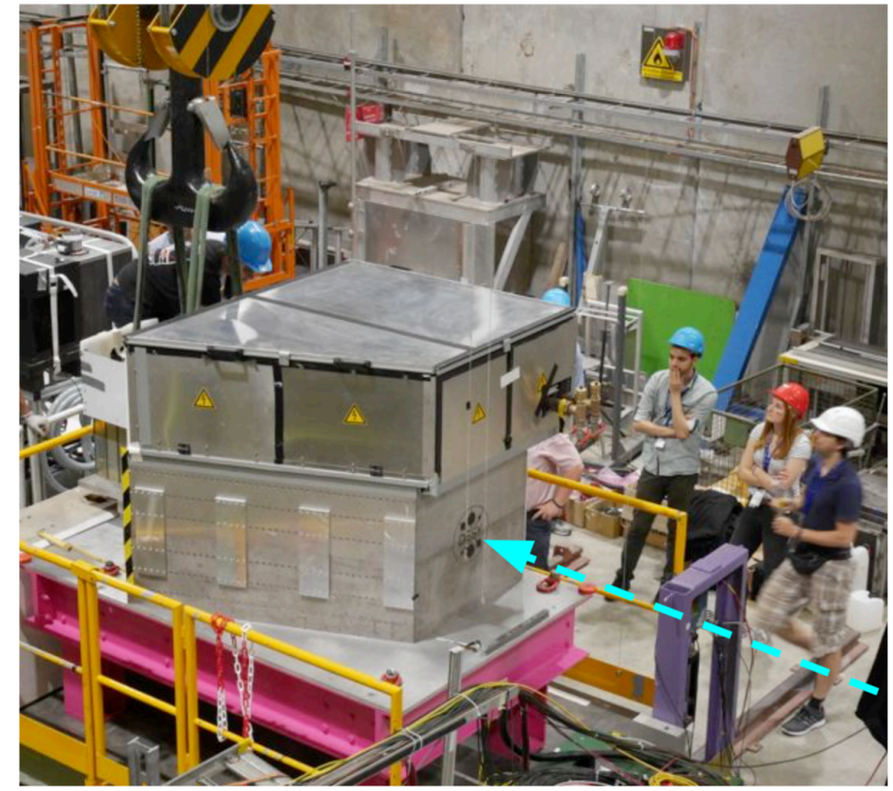
Weight Uncertainty in Neural Networks
C Blundell, J Cornebise, K Kavukcuoglu, D Wierstra
1505.05424

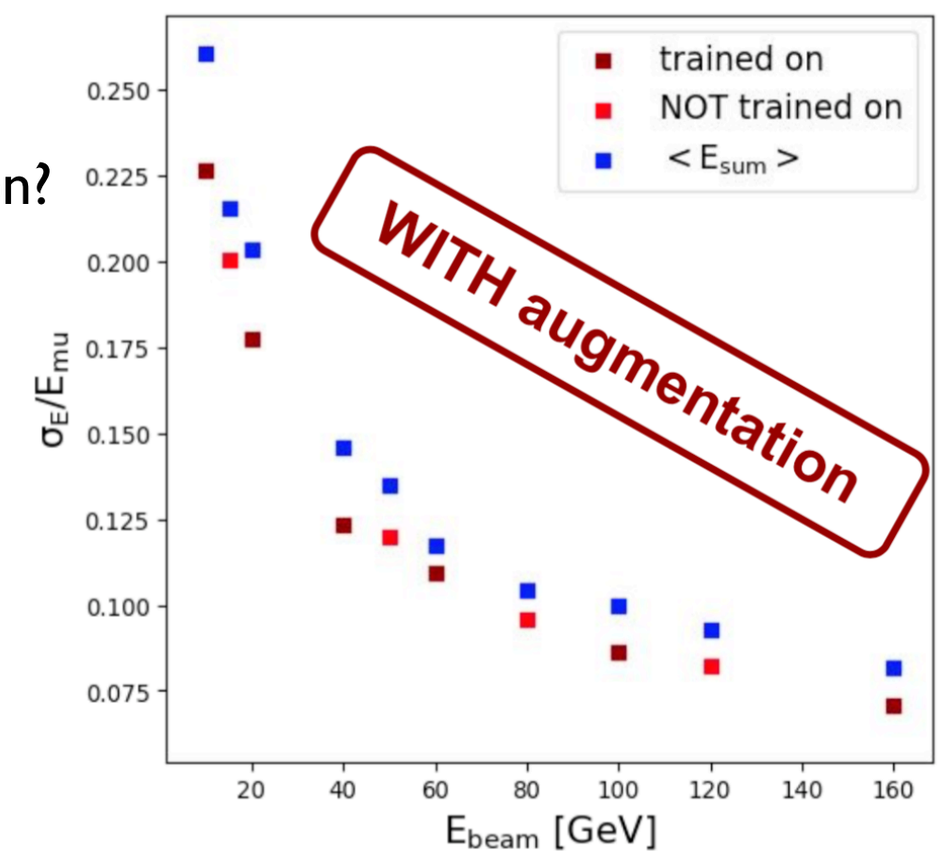
Plot by Sven Bollweg (BSc thesis), work with Heidelberg (Plehn),

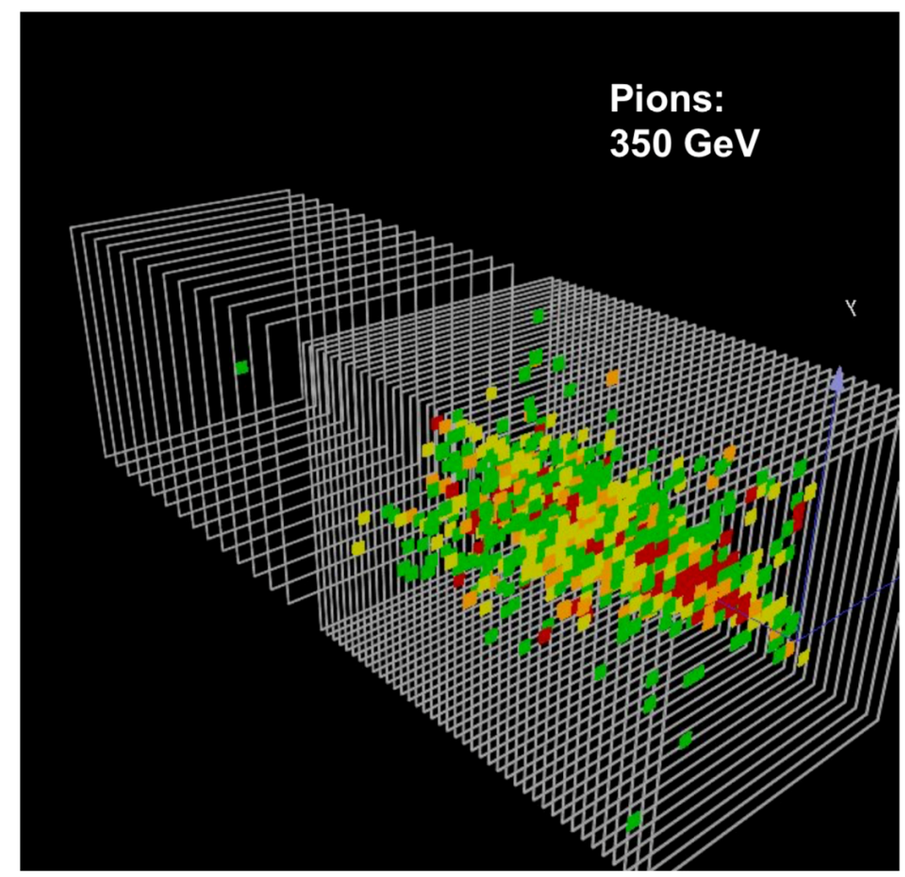
Calorimetry

MSc Student E. Buhmann, with E. Garutti

- Studies of hadronic calorimetry for a future e+e- collider (CALICE prototype)
- Technological prototype
  - 24*24*38 = ~22k tiles
  - 4 interaction lengths (steel as passive material)
- Study energy reconstruction/particle ID/shower separation using CNN/LCNs
  - ADC to GeV?
  - 5D Reconstruction?





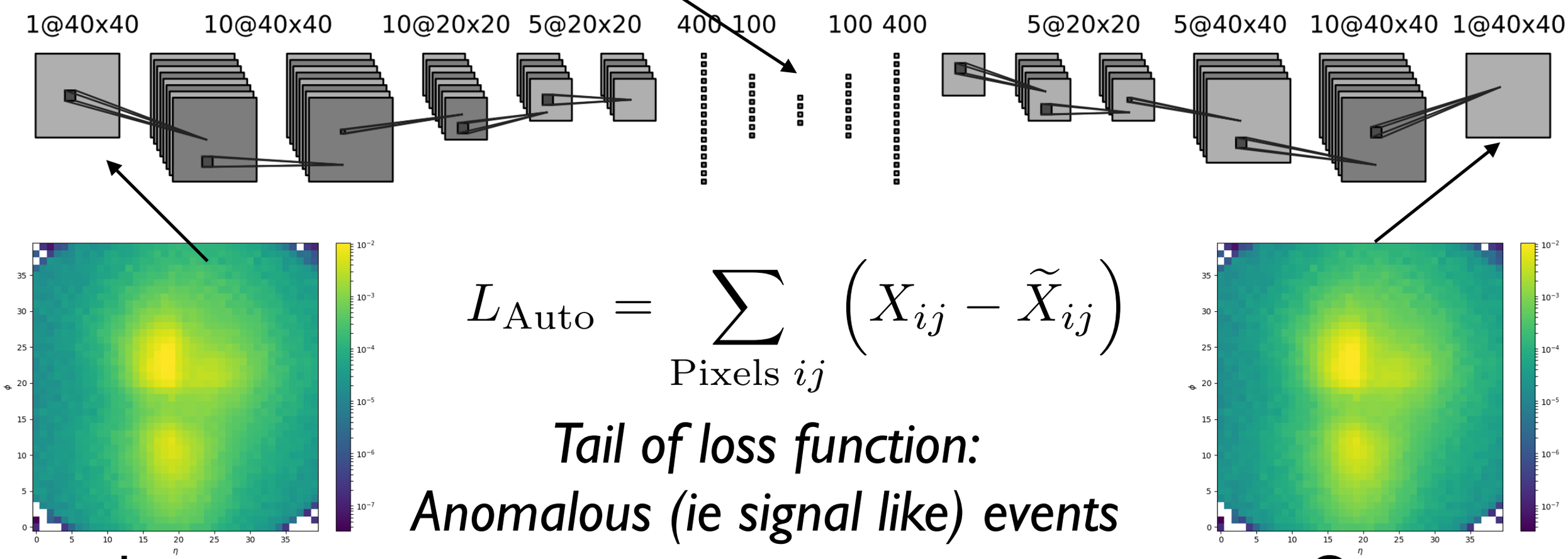


Related Topic: DNNs on FPGAs

Model independent searches

- Orthogonal to existing approaches:
  - Search for new physics without assuming a specific model
- Use (variational/adversarial) autoencoder trained on **data**
  - Detect new physics as anomaly
  - Not affected by usual systematic uncertainties
- Inclusion in trigger
- What can we learn from latent space?

QCD or What?
T Heimel, GK, T Plehn, JM Thompson, 1808.08979
Searching for New Physics with Deep Autoencoders
M Farina, Y Nakai, David Shih, 1808.08992



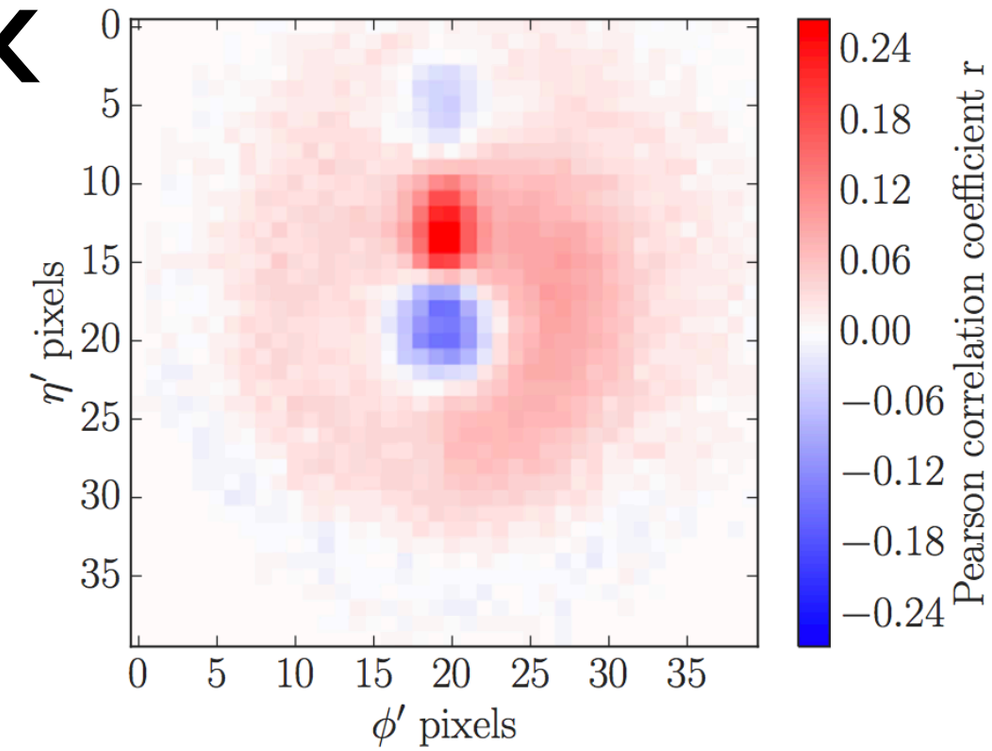
$$L_{\text{Auto}} = \sum_{\text{Pixels } ij} \left(X_{ij} - \widetilde{X}_{ij} \right)$$

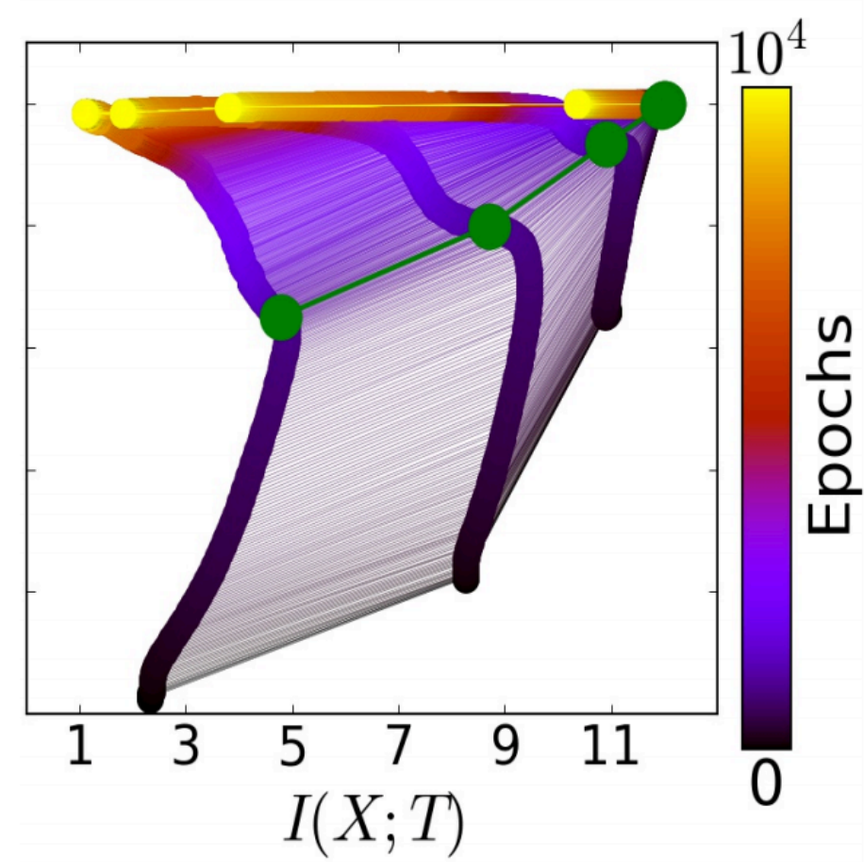
Tail of loss function:
Anomalous (ie signal like) events

Work with Heidelberg (Plehn)

Opening the Black Box

- What is necessary to trust a new algorithm?
  - **Learn what the network learns!**
    - Visualise decision process
    - Correlations & Gradients
    - Encode physics in the network
    - External known variables (special case in physics)
    - Interpretation of capsules?
  - Information theoretic approaches
  - Latent space of auto encoders





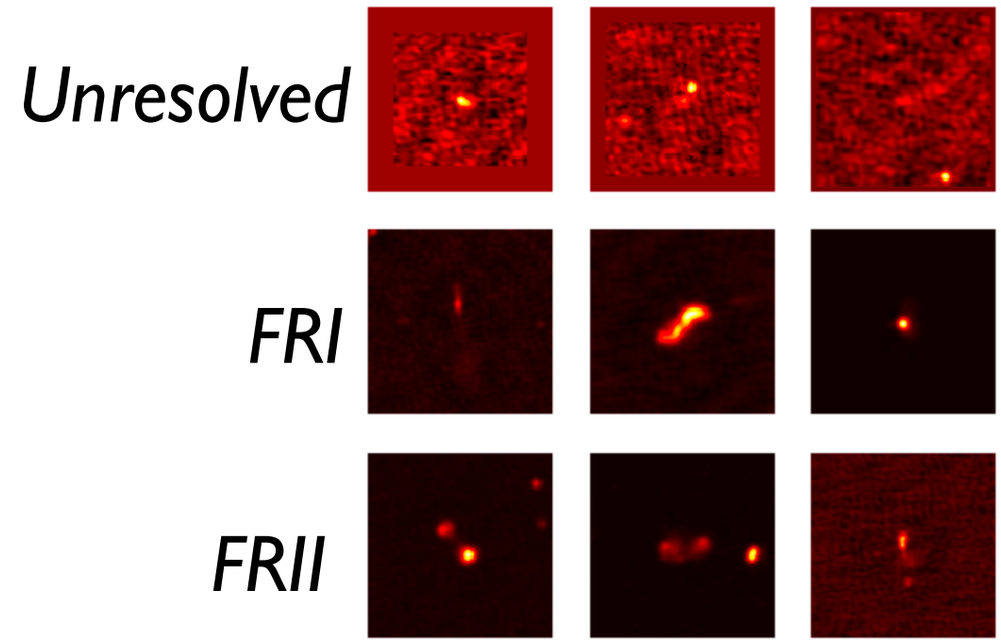
On the Information Bottleneck Theory of Deep Learning, Saxe et al, ICLR Proc 2018
Opening the Black Box of Deep Neural Networks via Information
Ravid Shwartz-Ziv, Naftali Tishby
1703.00810

Work with Heidelberg (Plehn)

New Connections

with V. Lukic and M. Brueggen

- Classification of radio galaxies
- Consider three morphologies
  - Unresolved
  - FRI (core-bright)
  - FRII (edge-bright)
- LOFAR Two-metre Sky Survey (LoTSS) at 150 MHz
  - Total of ~300k sources
  - Remaining ~3k after selecting (expert) classified images of active galactic nuclei
- Classify using DNN:
  - Convolutional and Capsule architectures



Transfer Learning

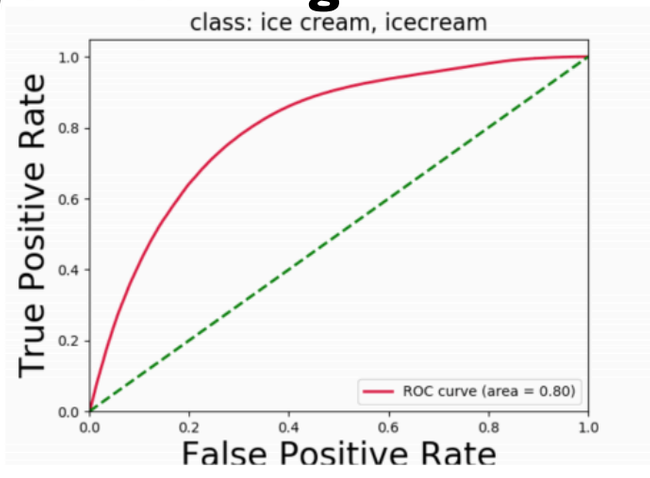
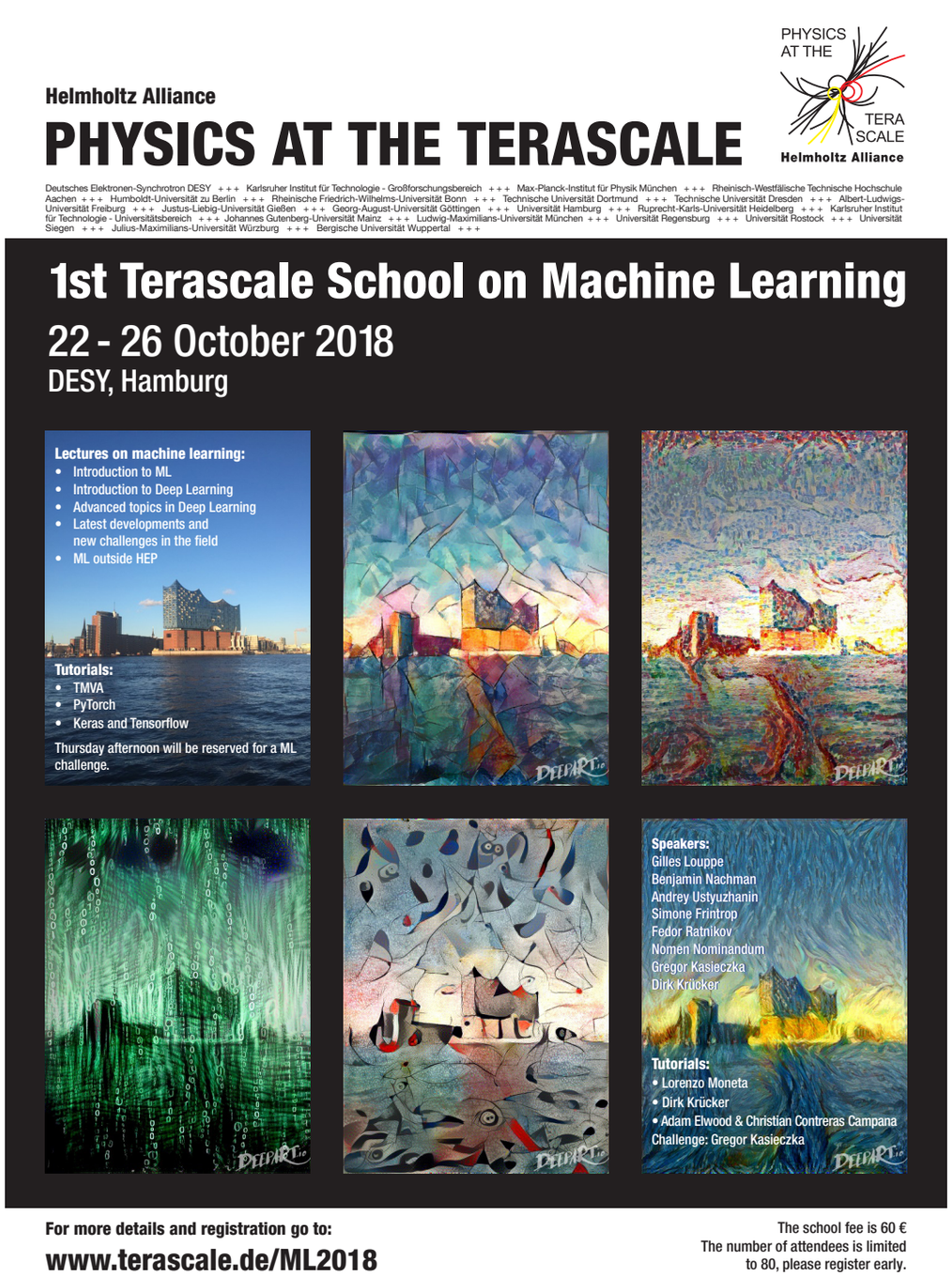
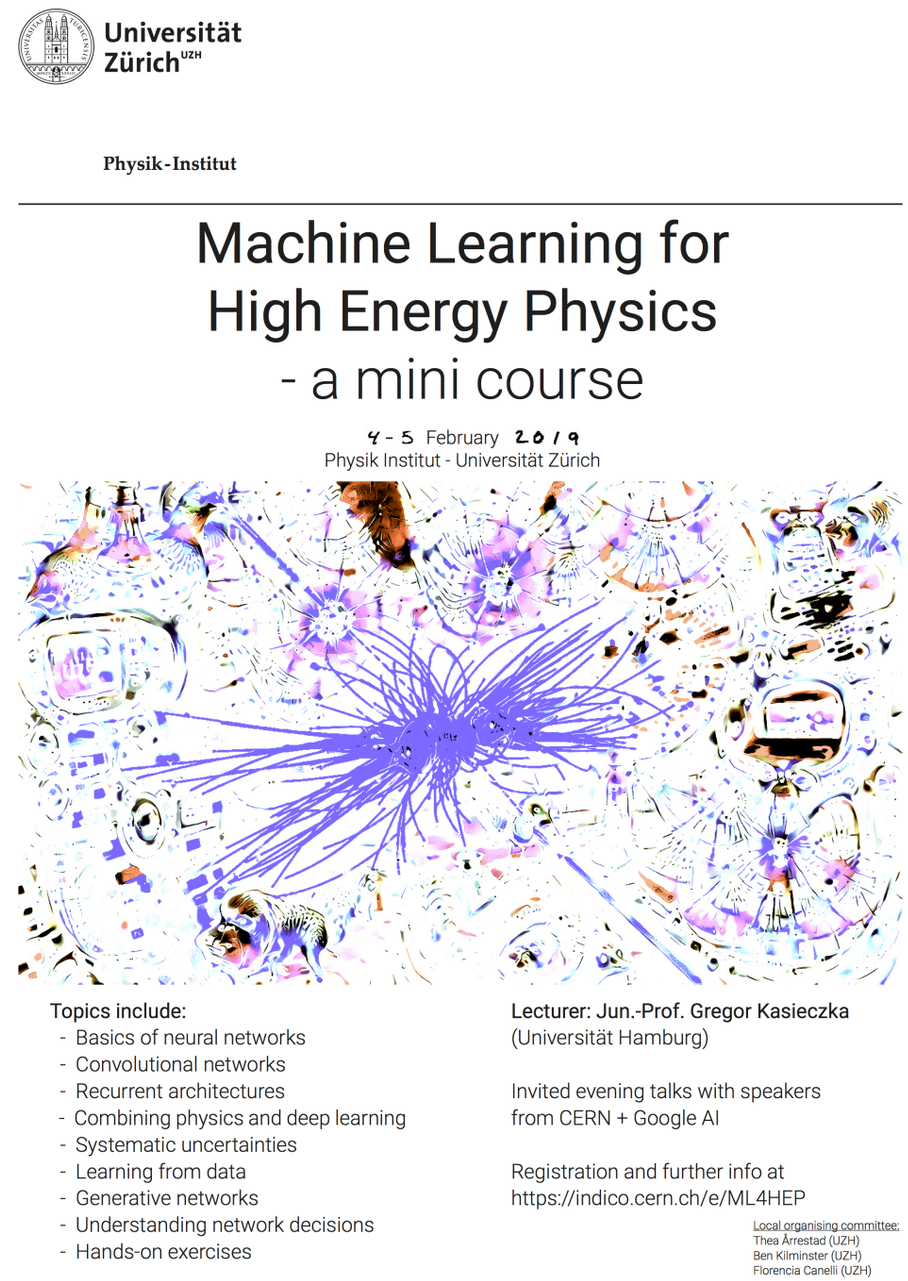


Image representation of data allows new shared algorithms across fields (astro, biology,…)



Education



- This year in cooperation with Yandex
- July 1-10 2019
- https://indico.cern.ch/event/768915/

…as well as integration in BSc

Integration

Cluster of Excellence: Quantum Universe
Platform for Challenges in Data Science

DASHH Data Science in Hamburg HELMHOLTZ Graduate School for the Structure of Matter

Deep Learning in the Natural Sciences
with emphasis on Astronomy, Particle Physics and Photonics

When? 28. 2. - 1. 3. 2019
Where? University of Hamburg
Jungiusstraße 9 HS III

Register here: https://indico.cern.ch/e/DLHAM

Invited speakers:
Dominik Elsässer
Martin Erdmann
Claudio Gheller
Filipe Maia
Tilman Plehn
Jesse Thaler

Organisers:
Gregor Kasieczka
Peter Schleper
Henry Chapman
Marcus Brüggen

PIER Partnership of Universität Hamburg and DESY
Universität Hamburg DER FORSCHUNG | DER LEHRE | DER BILDUNG
DESY.

Next Week

Hardware

- Currently included in DESY HPC Cluster Maxwell
  - O(10) nodes, each equipped with one NVidia Tesla P100
  - Shared between UHH and DESY CMS groups
- Considering moving to NAF:
  - Easier access to “our” data and software:
    - CVMFS, DUST, dCache mount
  - Testing GPUs already exist in NAF

Conclusions

- Deep learning and its applications to HEP is a lively and exciting research area
- Images are a powerful tool to represent physics/detector information (although they are not perfect)
  - Allow new kinds of collaborations
- Lots of progress in using networks for tasks beyond classification/regression
  - First steps towards asking qualitatively new kinds of questions
  - Low level reconstruction
- Better understanding of network decisions

Thank you!