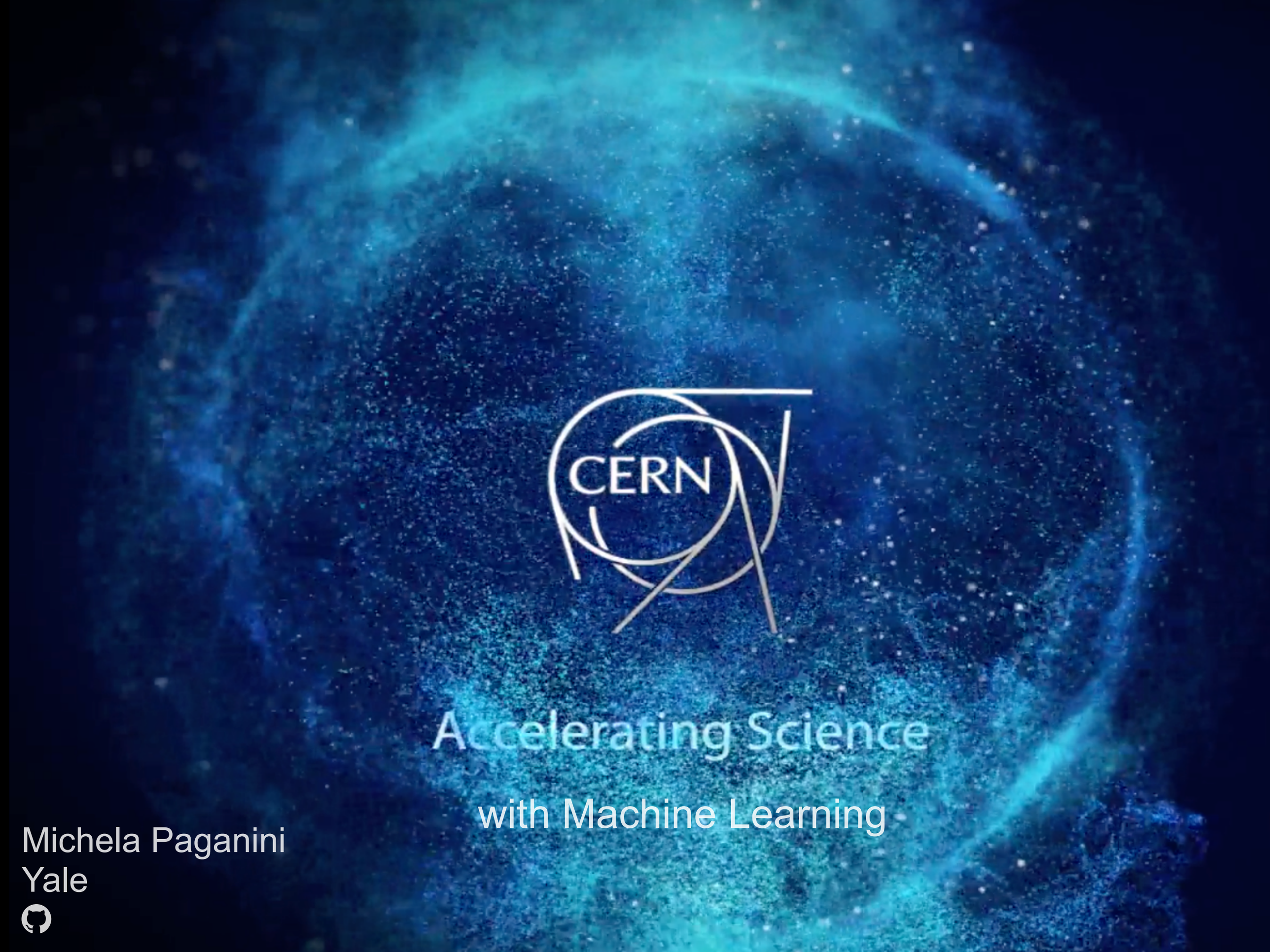
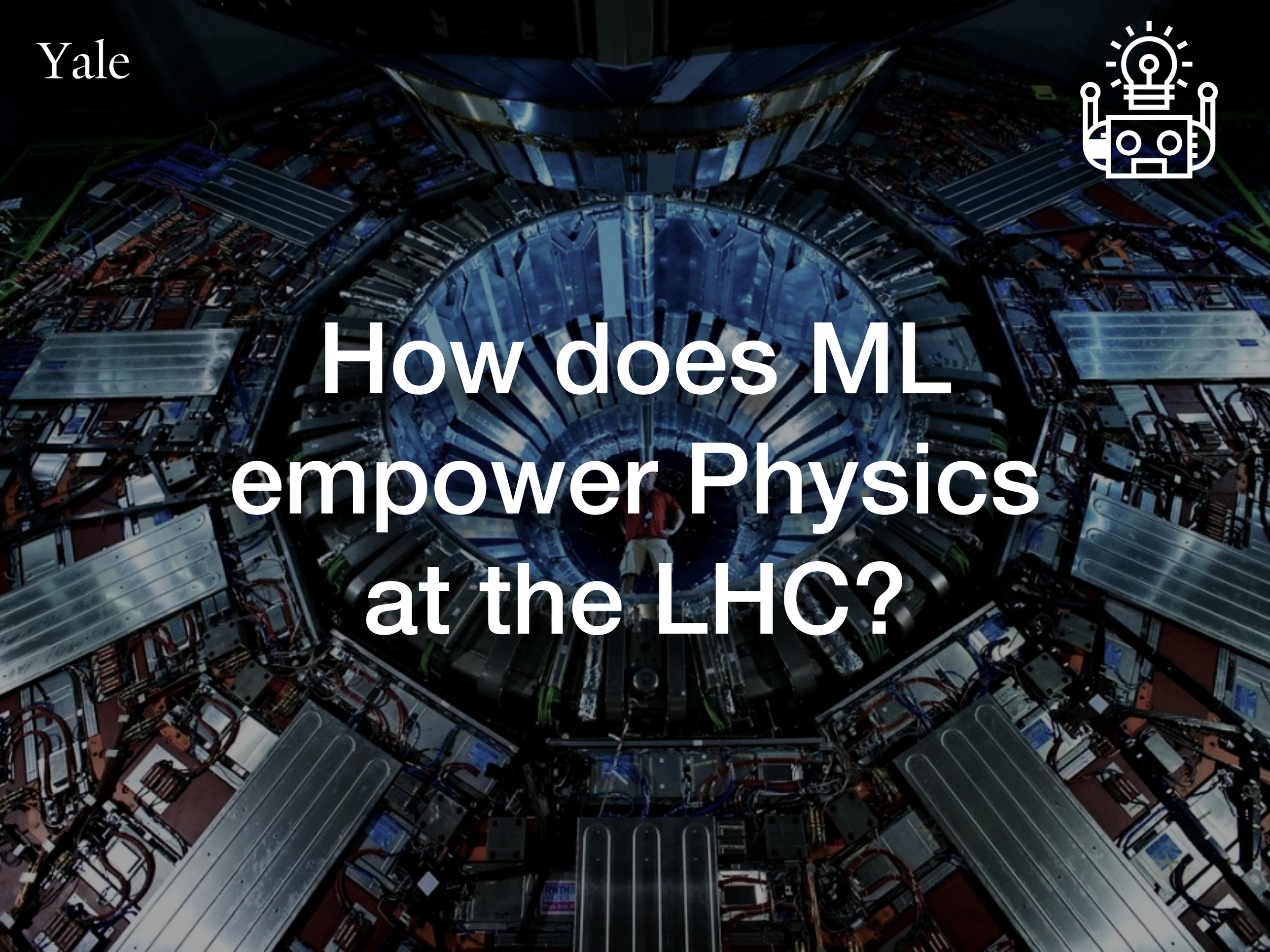


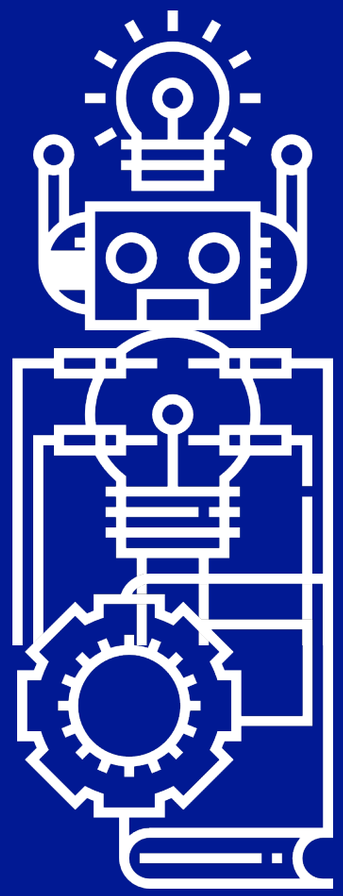
Michela Paganini
Yale

Yale



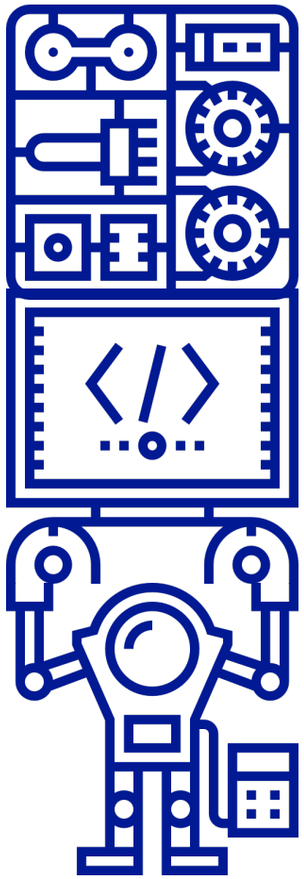
How does ML empower Physics at the LHC?

Yale



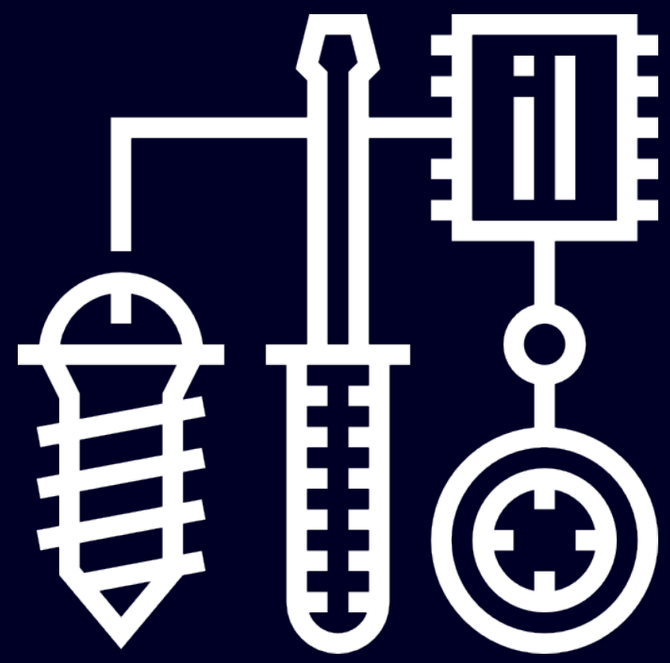
Just like integration has helped theorists

$$\frac{1}{AB} = \int_0^1 \mathrm{d}x \frac{1}{(xA + (1-x)B)^2} = \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{\delta(x+y-1)}{(xA+yB)^2}$$

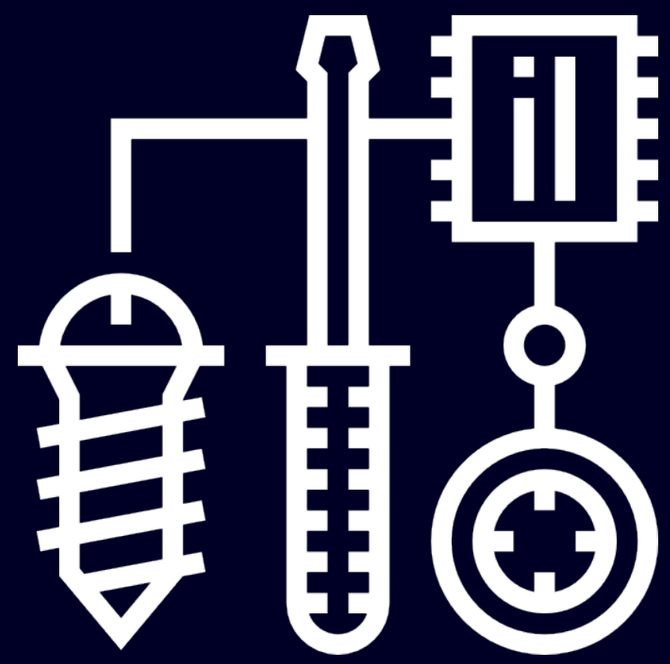


Just like statistics has helped experimentalists

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$



It’s a tool!

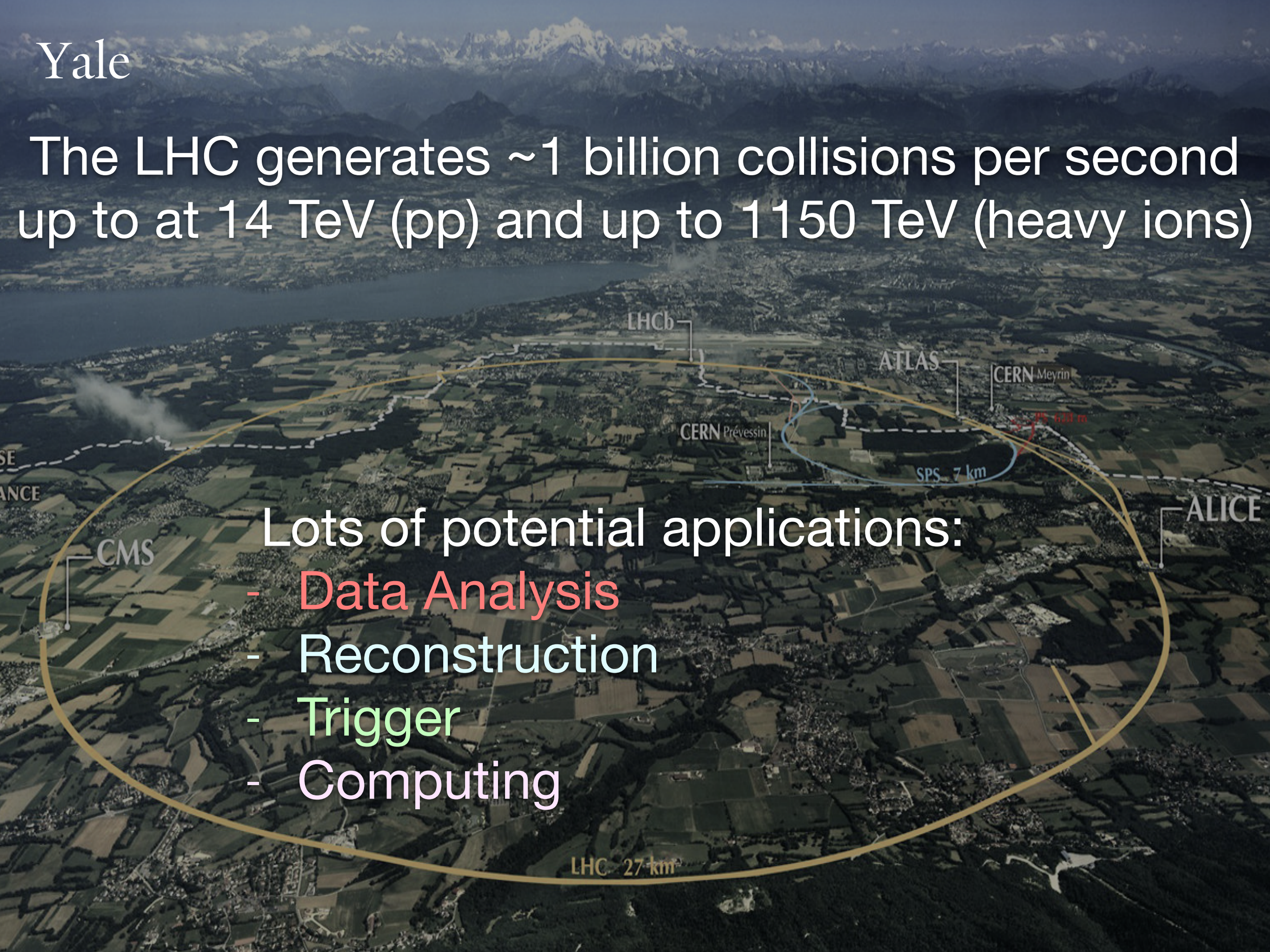


Well…
Like mathematics and statistics, machine learning is an **entire field of research**, both theoretical and experimental.
But for us at CERN, for all intents and purposes,

It’s a tool!

Yale

The LHC generates ~1 billion collisions per second up to at 14 TeV (pp) and up to 1150 TeV (heavy ions)

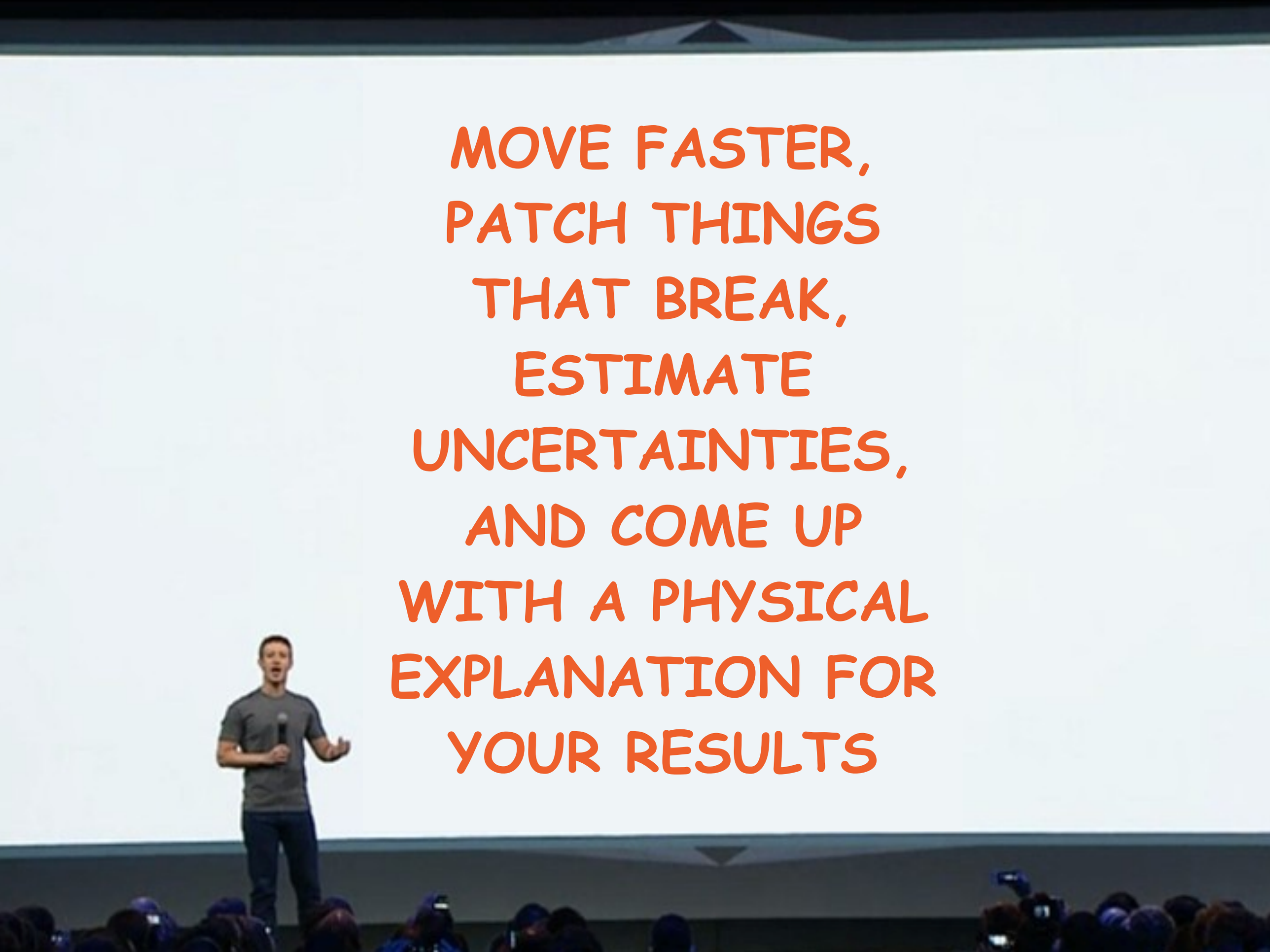


Lots of potential applications:

- Data Analysis
- Reconstruction
- Trigger
- Computing

MOVE
FAST AND
BREAK
THINGS

MOVE
FAST AND
BREAK
THINGS


MOVE FASTER,
PATCH THINGS
THAT BREAK,
ESTIMATE
UNCERTAINTIES,
AND COME UP
WITH A PHYSICAL
EXPLANATION FOR
YOUR RESULTS

Important Considerations for ML in HEP

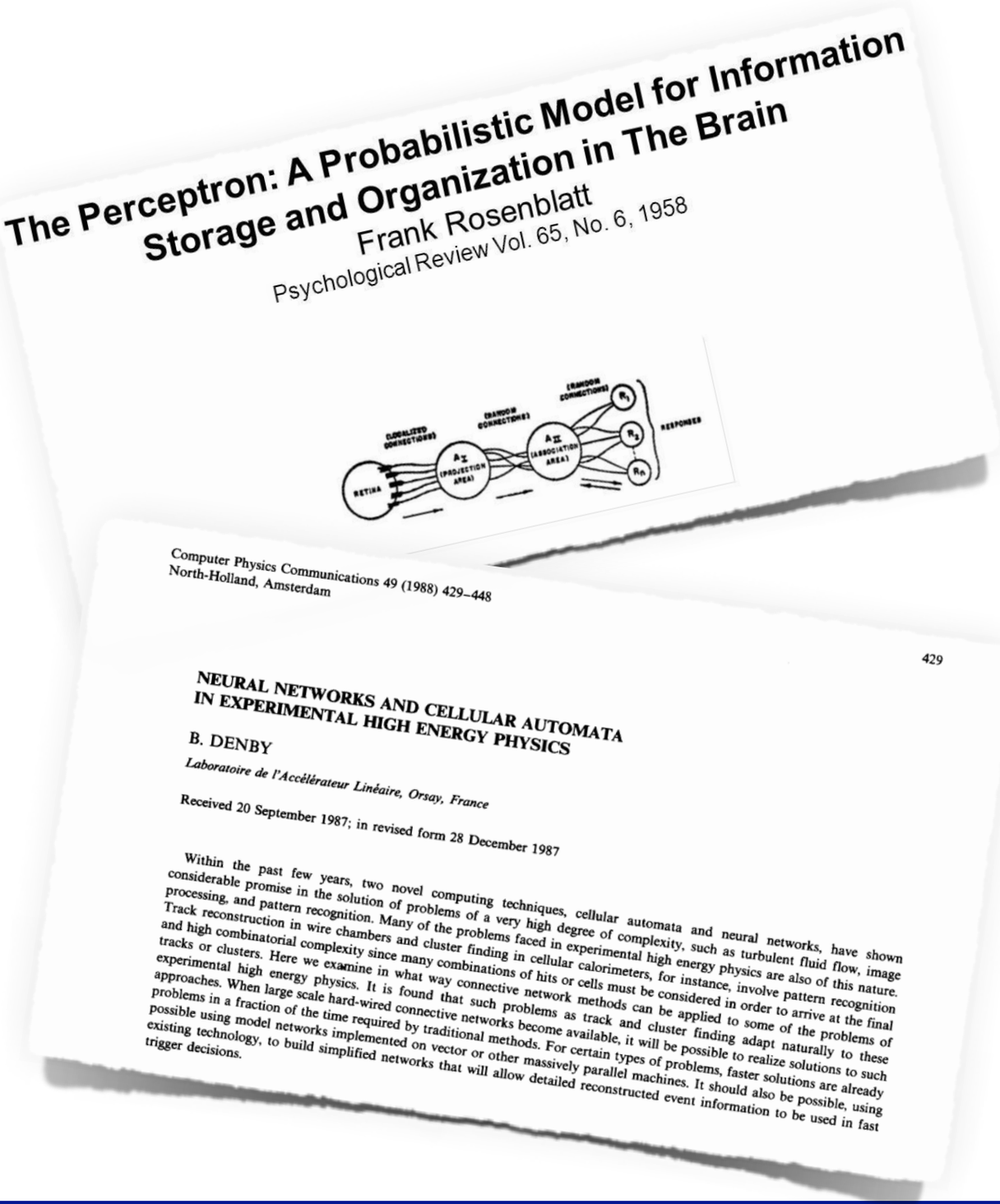
- **Performance:**
  - improve upon traditional methods
  - save time and/or CPU

- **Domain Specific Knowledge:**
  - insert Physics knowledge to guide or constrain the net

- **Interpretation:**
  - favor simpler models
  - investigate what's being learned and how throughout the layers
  - visualize inner activations and information propagation
  - learn from machine learning: helpful to design new physics variables

- **Confidence:**
  - need to estimate uncertainties

Look how far we have come!



The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain

Frank Rosenblatt

Psychological Review Vol. 65, No. 6, 1958

Computer Physics Communications 49 (1988) 429–448
North-Holland, Amsterdam

429

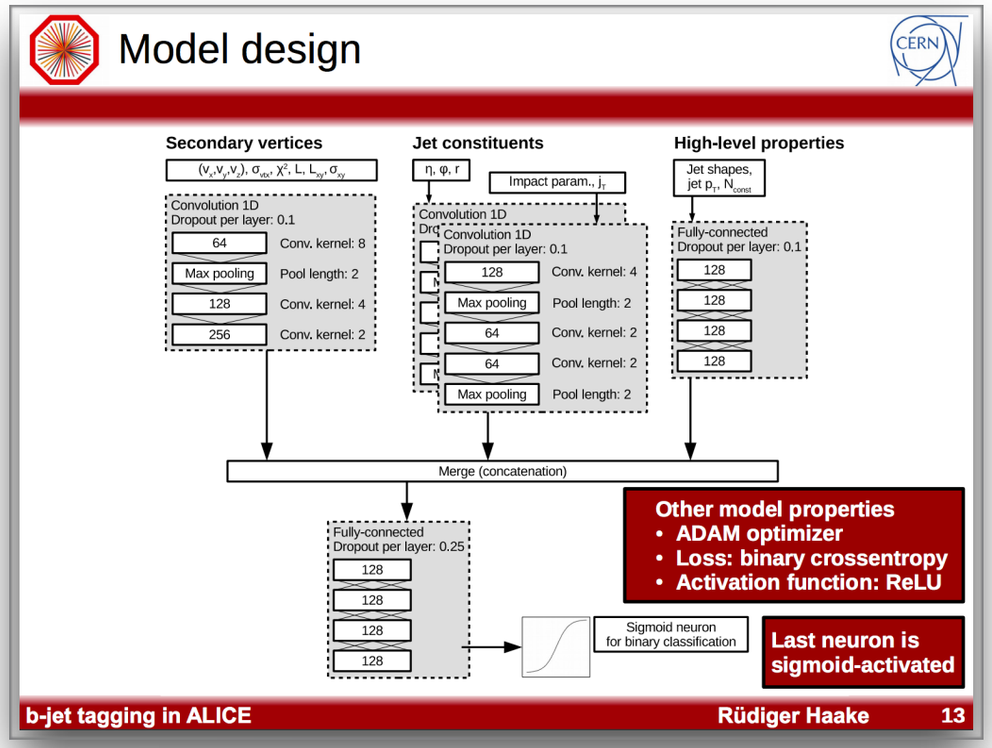
NEURAL NETWORKS AND CELLULAR AUTOMATA IN EXPERIMENTAL HIGH ENERGY PHYSICS

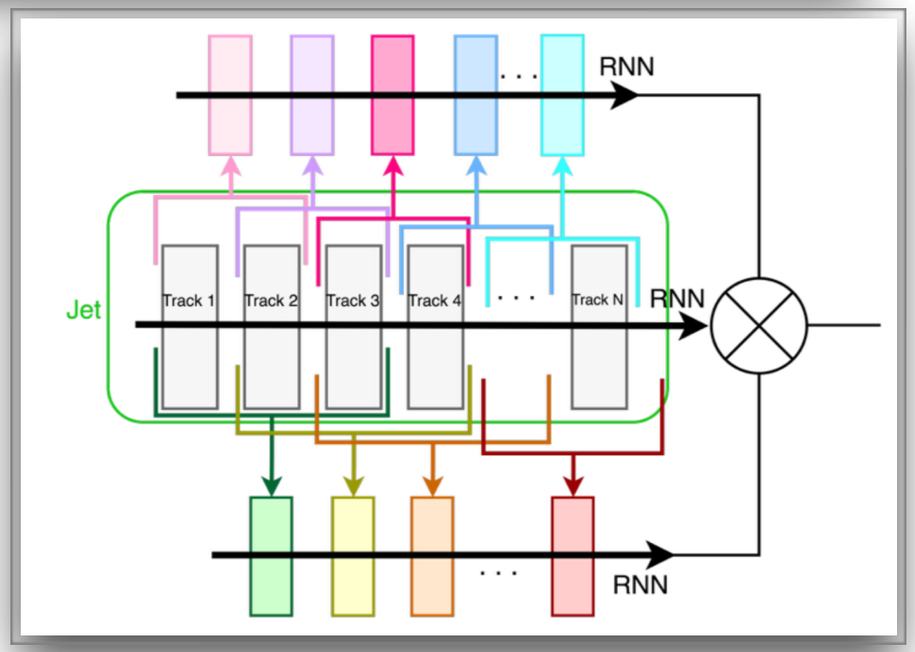
B. DENBY

Laboratoire de l'Accélérateur Linéaire, Orsay, France

Received 20 September 1987; in revised form 28 December 1987

Within the past few years, two novel computing techniques, cellular automata and neural networks, have shown considerable promise in the solution of problems of a very high degree of complexity, such as turbulent fluid flow, image processing, and pattern recognition. Many of the problems faced in experimental high energy physics are also of this nature. Track reconstruction in wire chambers and cluster finding in cellular calorimeters, for instance, involve pattern recognition and high combinatorial complexity since many combinations of hits or cells must be considered in order to arrive at the final tracks or clusters. Here we examine in what way connective network methods can be applied to some of the problems of experimental high energy physics. It is found that such problems as track and cluster finding adapt naturally to these approaches. When large scale hard-wired connective networks become available, it will be possible to realize solutions to such problems in a fraction of the time required by traditional methods. For certain types of problems, faster solutions are already possible using model networks implemented on vector or other massively parallel machines. It should also be possible, using existing technology, to build simplified networks that will allow detailed reconstructed event information to be used in fast trigger decisions.

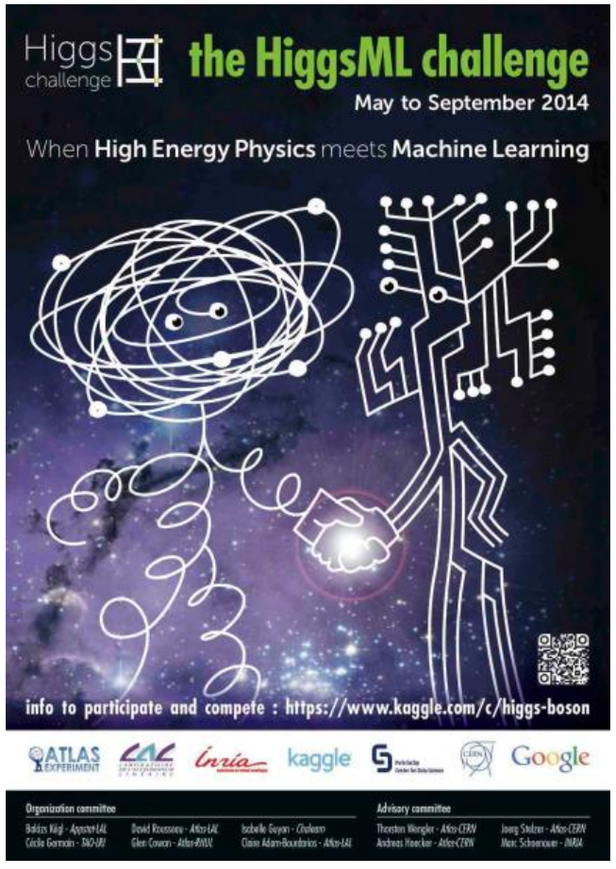




Yale

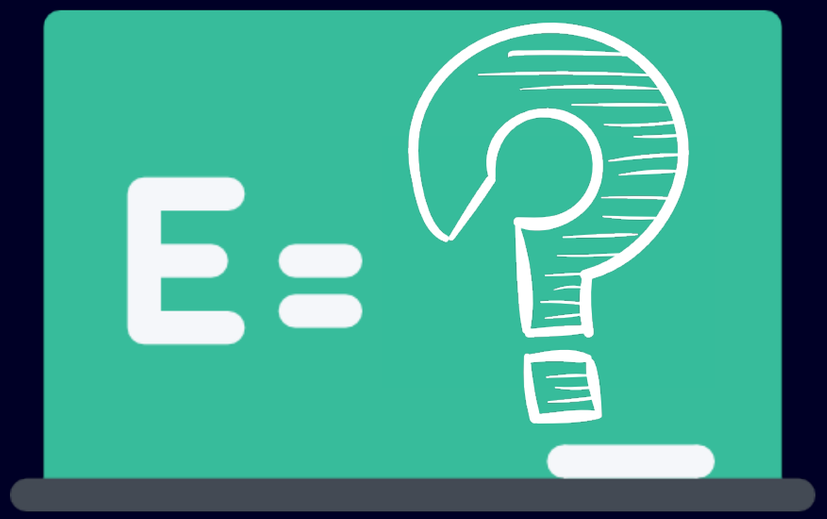
- We lack the software (or hardware) infrastructure for fast experimentation of state-of-the-art solutions

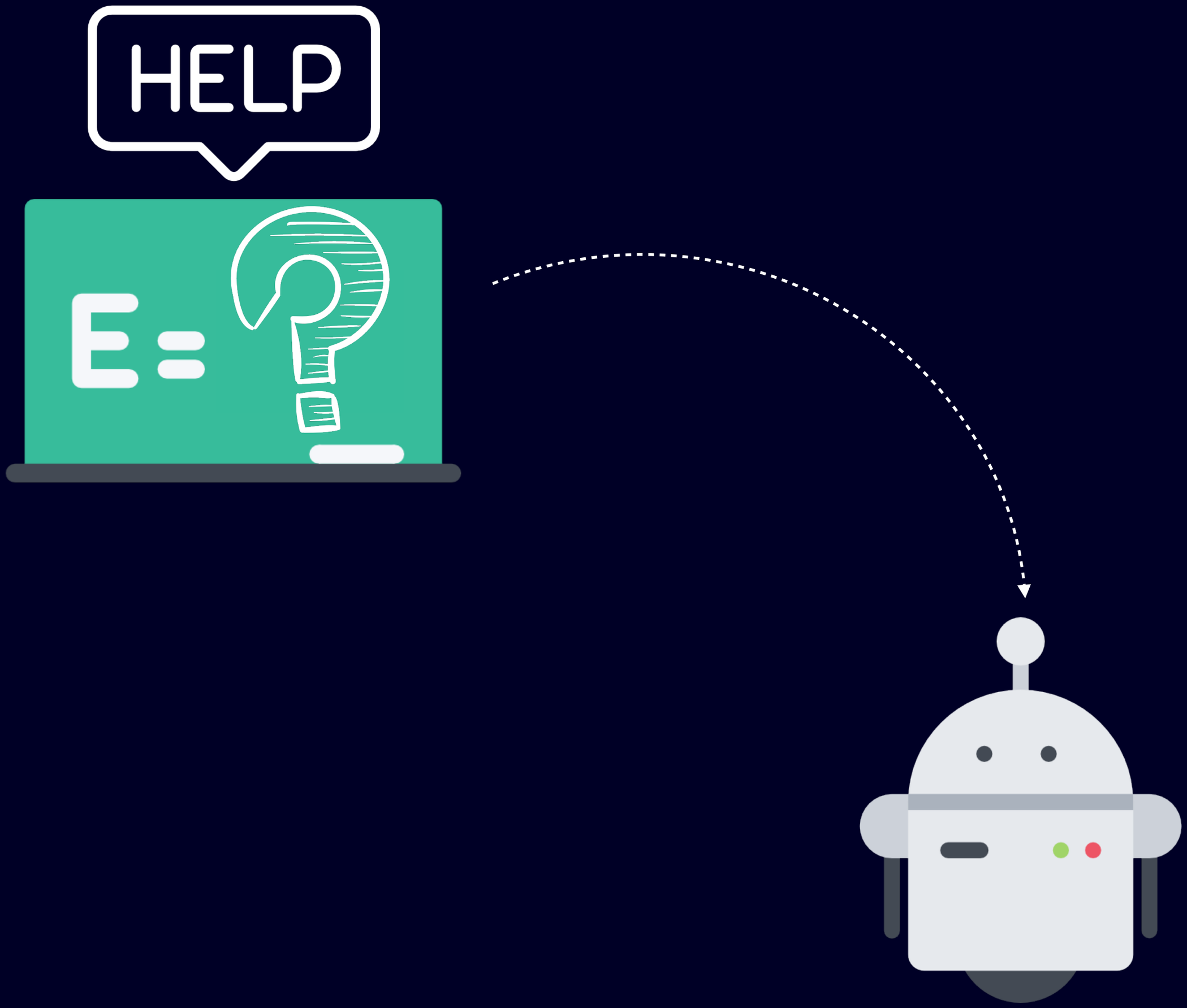
- Yet, growing challenges + new young talents have forced the various LHC experiments to slowly join the deep learning revolution

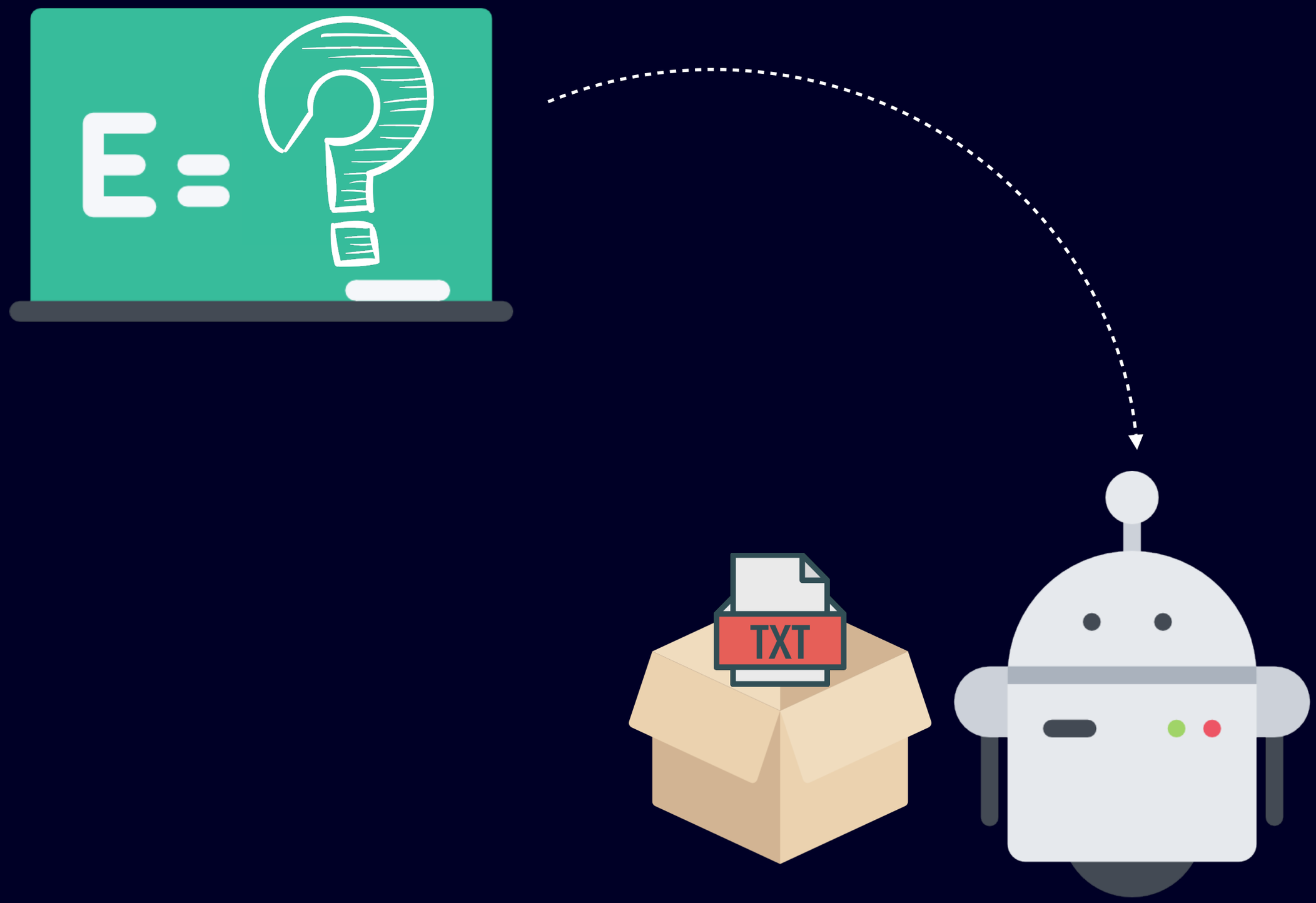


- By the time we started caring, DL had already exploded and created solutions that we can now borrow to solve our problems

E=

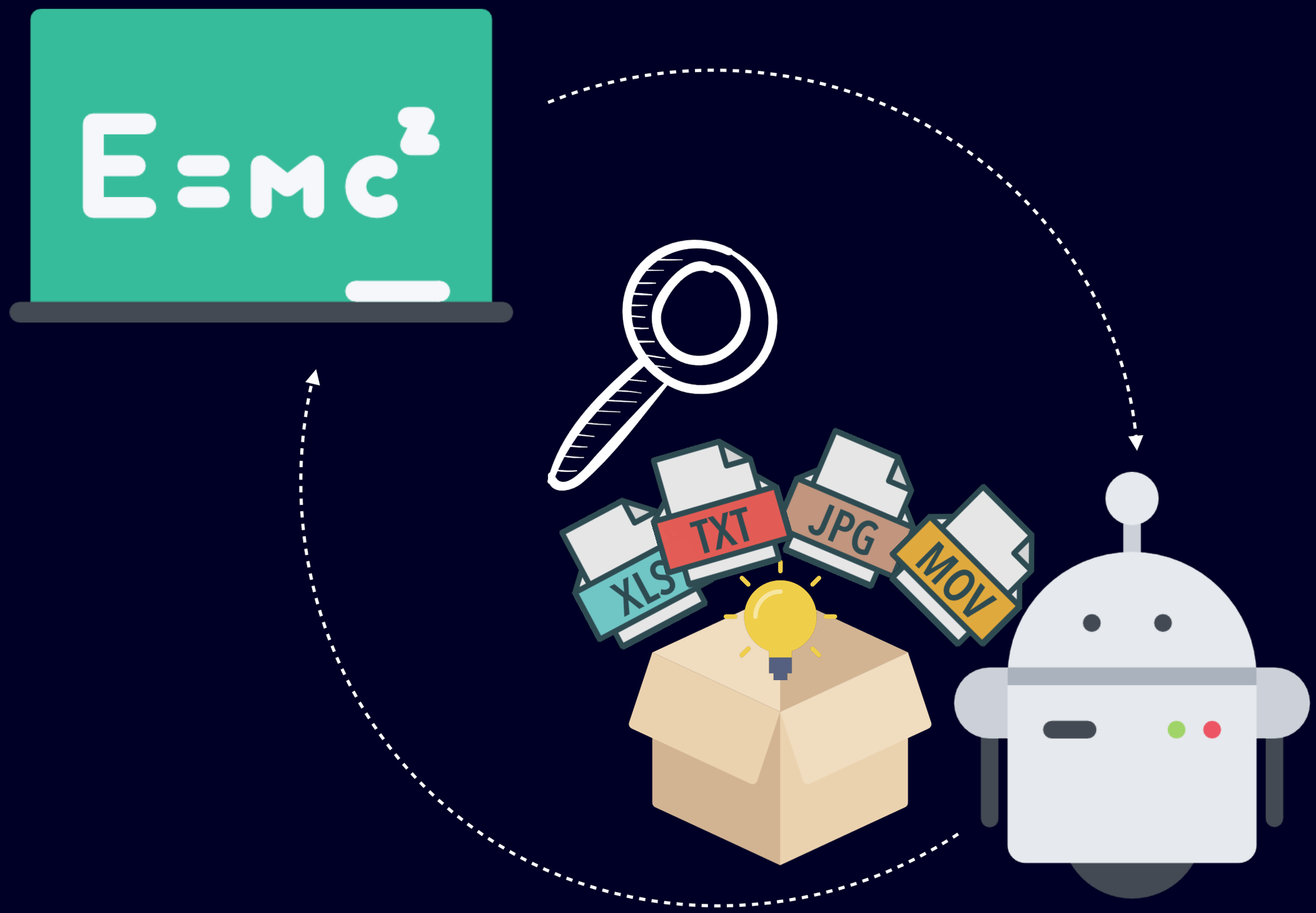

E=


HELP
E=


E=
TXT

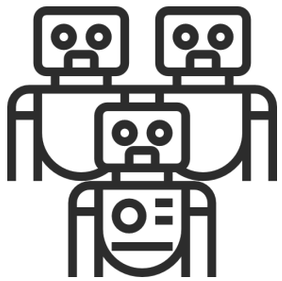
E=
XLS
TXT
JPG
MOV

E=
XLS
TXT
JPG
MOV


E=mc²
XLS
TXT
JPG
MOV

Yale

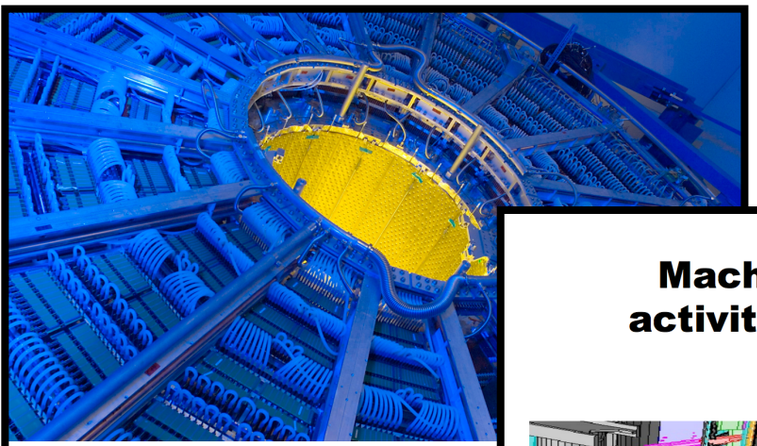
- In certain cases, this search might fail — either no similar problems, or no applicable solution



- New ML solutions from physicists have been pushing the boundaries of ML

Get the latest news from the CERN OpenLab workshop webcast!

ALICE



ALICE: ML and DA Challenges

ATLAS EXPERIMENT

Machine Learning in ATLAS: activities and future challenges

Matthew Beckingham, Michael Kagan, David Rousseau for the ATLAS collaboration

OpenLab ML and Analytics workshop, 29th April 2016

LHCb

LHCb Machine Learning challenges

, Stefan Roiser[1], Andrey Ustyuzhanin[1,2]

dex School of Data Analysis

CMS

Projects proposal at the CMS Experiment

CERN Openlab Machine Learning and Data Analytics Workshop
29 April 2016

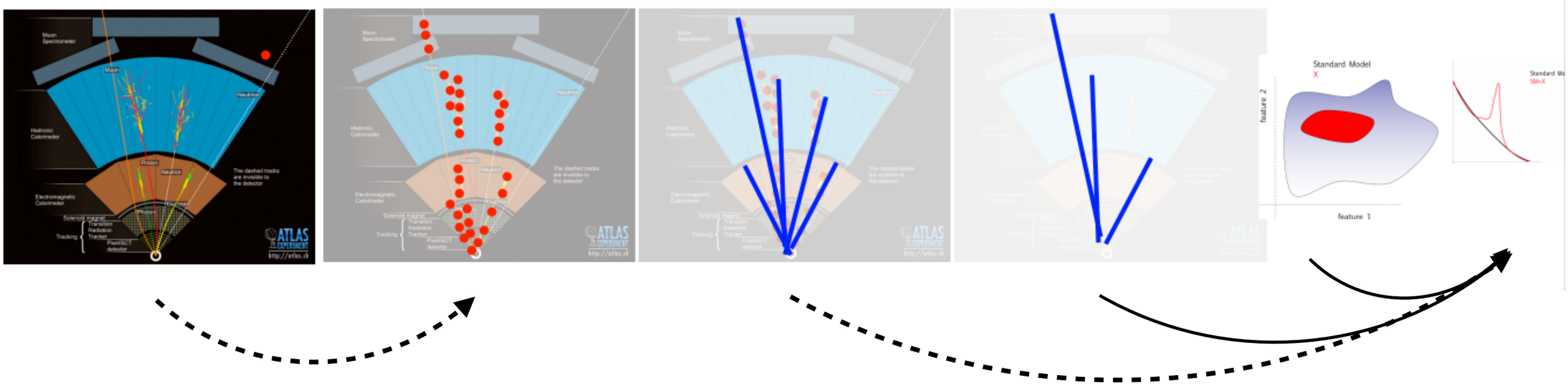
Recent Workshops and Conferences on ML in HEP

- NIPS Workshop 2014, 2015
- DS@HEP 2015, 2016, 2017
- Connecting the Dots 2016, 2017
- IML Workshop
- CERN OpenLab workshop on ML and Data Analytics

From Daniel Whiteson, NIPS 2015

Dimensionality

Raw	Sparsified	Reco	Select	Physics	Ana
1e7	1e4	100-ish*	50	10	1



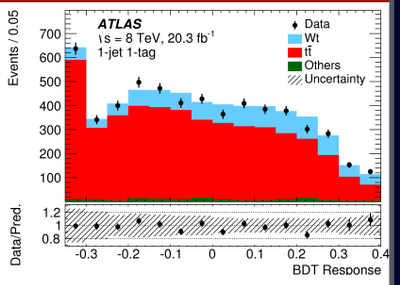
How far can we push it?

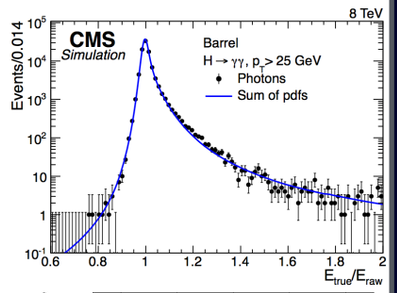
From Michael Kagan

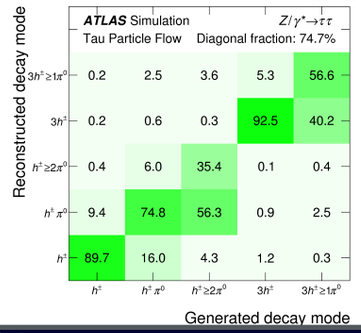
How Does Machine Learning Fit In?

15

- **In analysis**:
  - Classifying signal from background, especially in complex final states
  - Reconstructing heavy particles and improving the energy / mass resolution
- **In reconstruction**:
  - Improving detector level inputs to reconstruction
  - Particle identification tasks
  - Energy / direction calibration
- **In the trigger**:
  - Quickly identifying complex final states
- **In computing**:
  - Estimating dataset popularity, and determining how number and location of dataset replicas





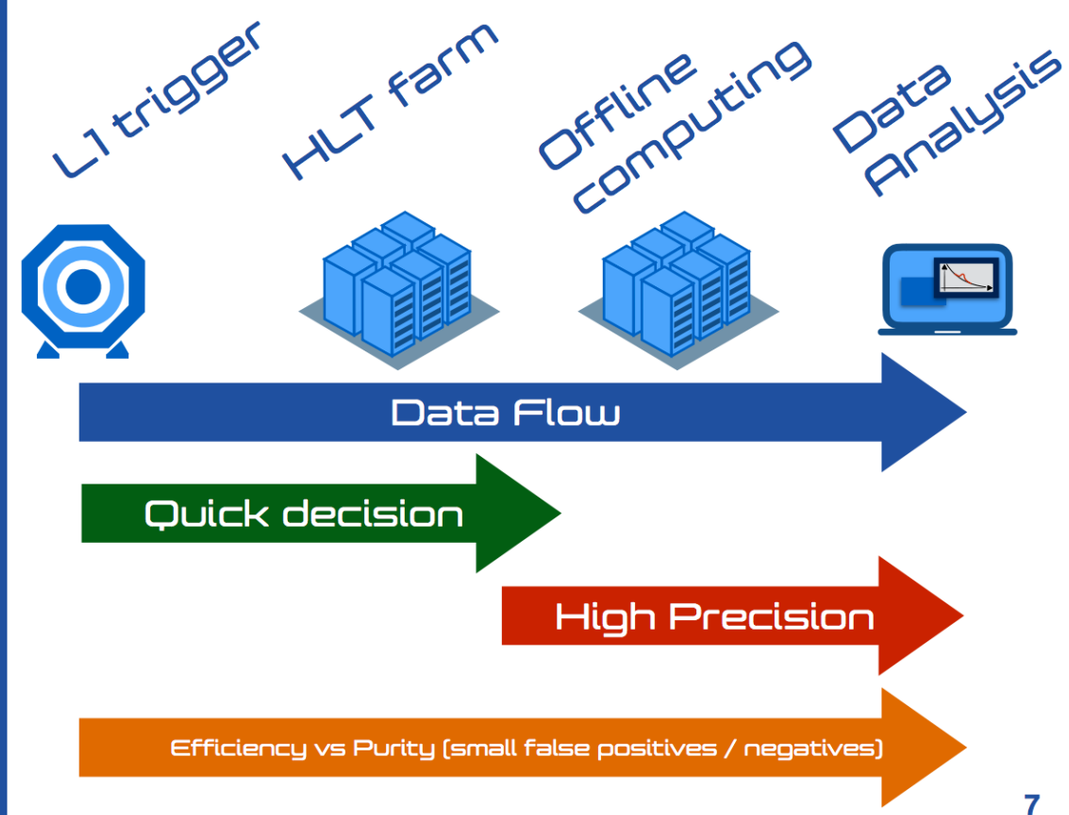


From Jean-Roch Vlimant

- Data taking
  - Real time event categorization
  - Data monitoring & certification robot
- Data Reconstruction
  - Calorimeter reconstruction
  - Boosted object jet tagging
- Data Processing
  - Computing Resource Optimization
  - Predicting data popularity
  - Intelligent networking
- Data Analysis
  - CMS assistance service
  - Big data reduction and analysis
  - Model independent search

From Maurizio Pierini

Four HEP environments

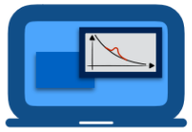


7

What DL can do for us

Online Data Taking (real time)
Fast trigger algorithms for topology classification
Fast reconstruction algorithms (clustering, tracking, classification)
Monitor detector operation conditions & data quality

Offline event processing (centralised)
Event indexing based on topology classification
Fast collision simulation based on generative models
Fast reconstruction algorithms (clustering, tracking, classification)



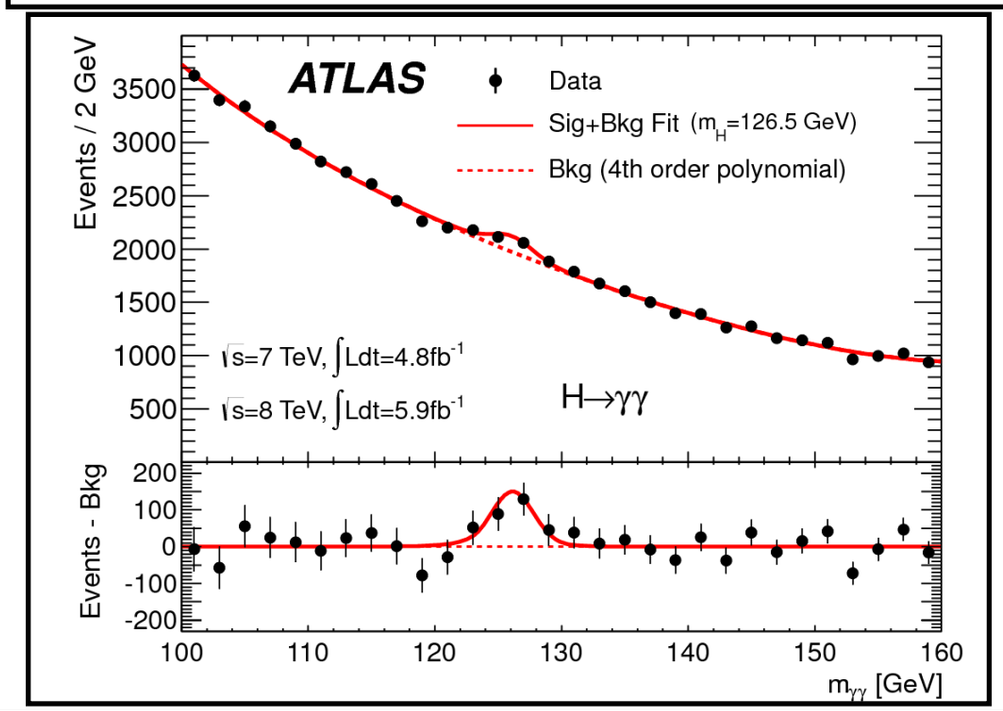
Data analysis (by users)
Particle identification
Jet tagging (g vs q vs b vs W/Z vs H vs t)
Unsupervised search for new physics as outlier detection

Analysis ++

(Kernel) Density Estimation

IML WG meeting: Unsupervised learning

Anton Poluektov



New approaches for boosting to uniformity

Alex Rogozhnikov[a,b*], Aleksandar Bukva[c], Vladimir Gligorov[d], Andrey Ustyuzhanin[b,e,f] and Mike Williams[g]

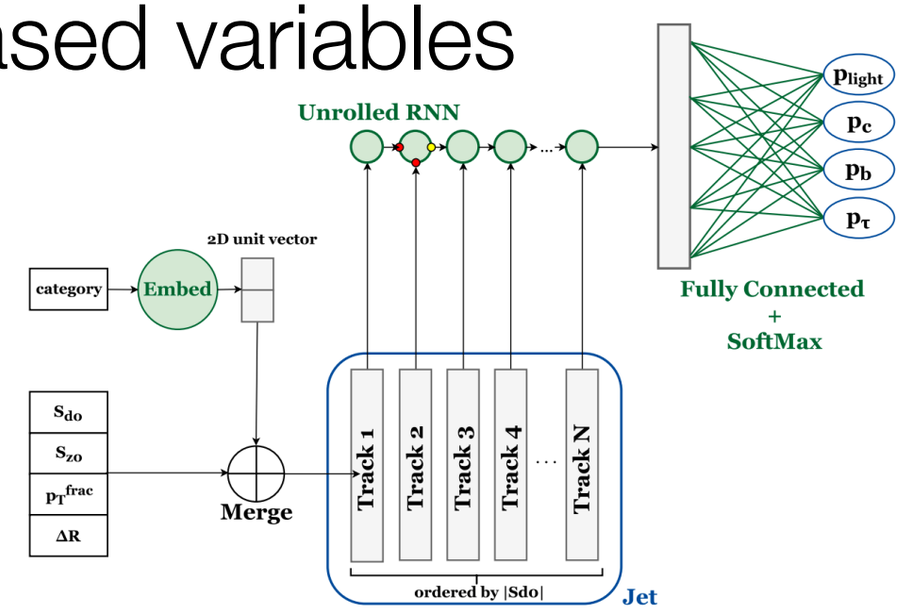
Neural Networks for the Extraction of the Λ_C Signal in p-Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV

Machine Learning Methods in the Analysis of Low-mass Dielectrons in ALICE

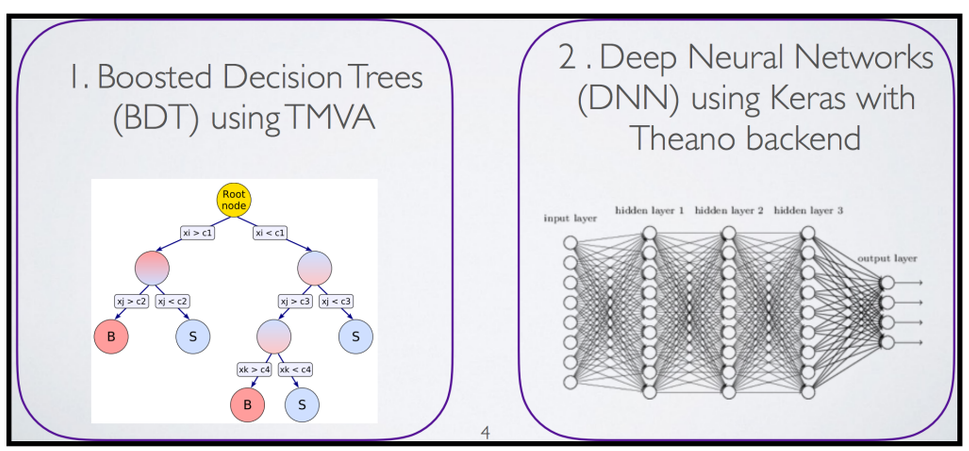
Sebastian Lehner and Aaron Capon**
on behalf of the ALICE collaboration

Jet Tagging

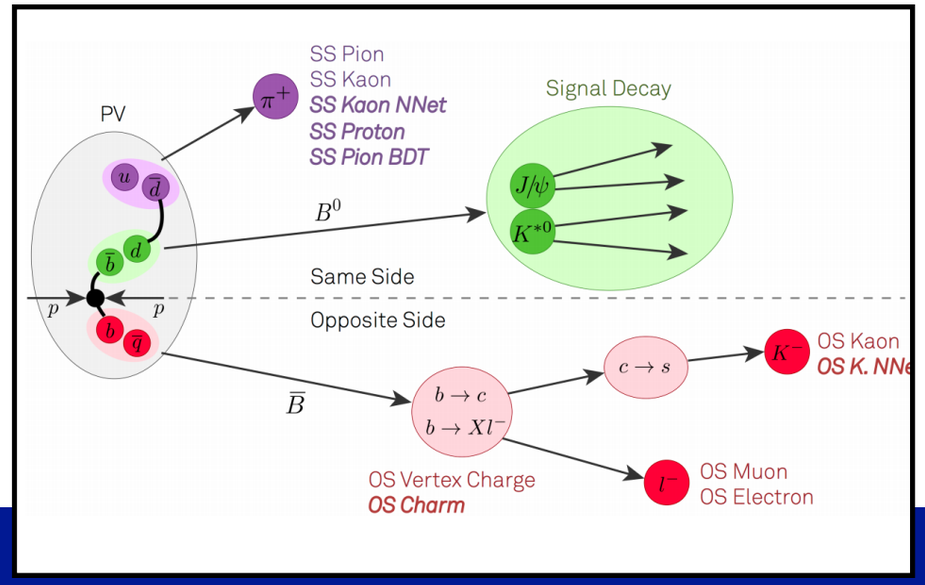
- Flavor tagging in ATLAS from track-based variables using Recurrent Neural Networks



- W boson and top tagging in ATLAS using high level features in BDTs and NNs



- Inclusive Flavor Tagging at LHCb

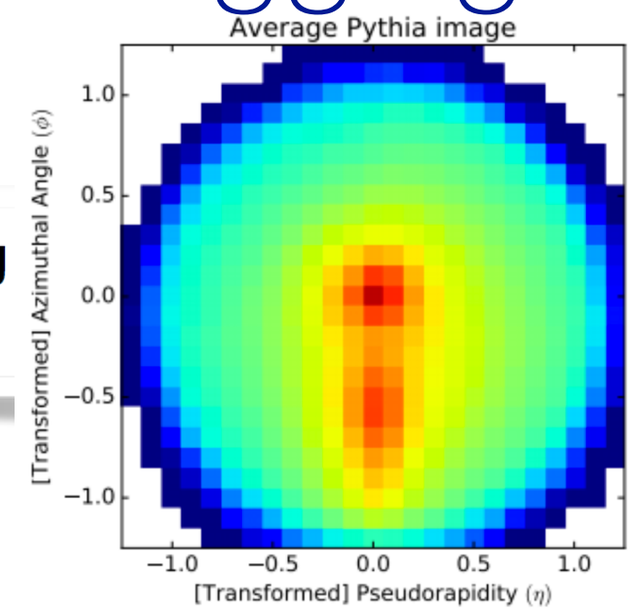


Computer Vision Solutions to Jet Tagging

- Based on *jet images*

Jet-Images: Computer Vision Inspired Techniques for Jet Tagging

Josh Cogan, Michael Kagan, Emanuel Strauss, Ariel Schwartzman



Other References:

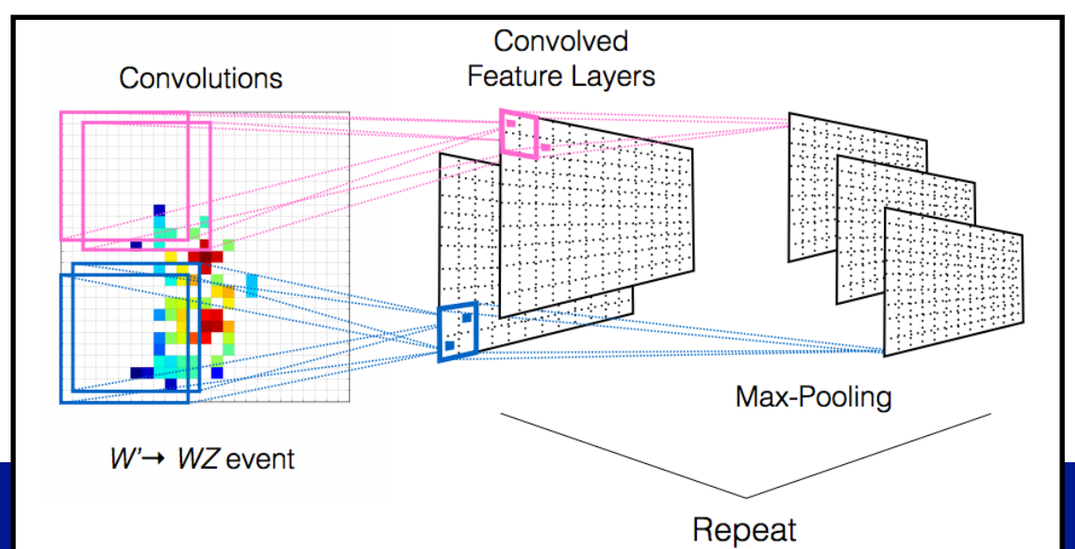
L. G. Almeida, M. Backovi´c, M. Cliche, S. J. Lee and M. Perelstein, *Playing Tag with ANN: Boosted Top Identification with Pattern Recognition*, JHEP 07 (2015) 086 [1501.05968].

P. T. Komiske, E. M. Metodiev and M. D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, [1612.01551].

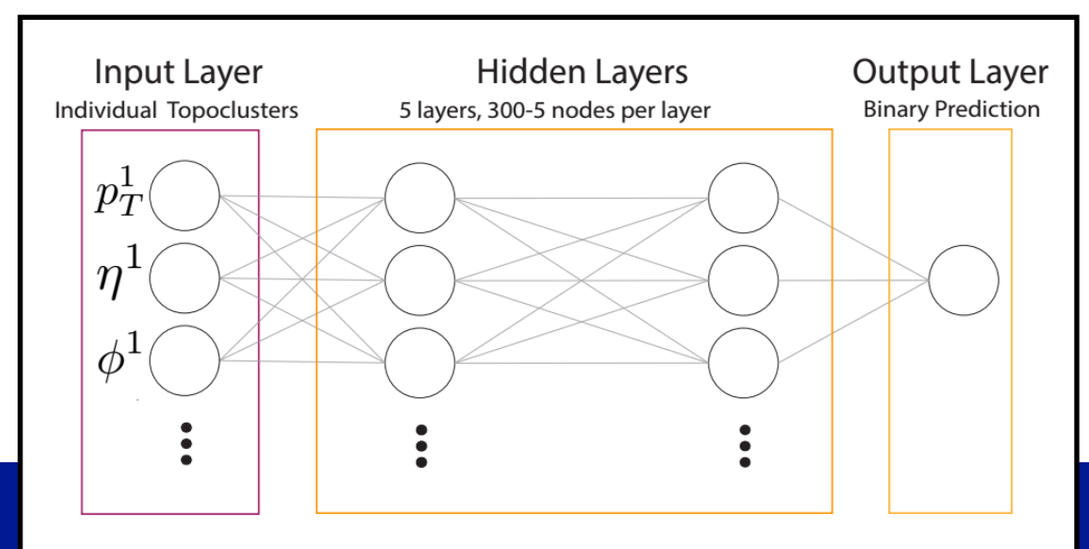
J. Barnard, E. N. Dawe, M. J. Dolan and N. Rajcic, *Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks,* [1609.00607].

P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet Substructure Classification in High-Energy Physics with Deep Neural Networks*, Phys. Rev. D93 (2016), no. 9 094034 [1603.09349

- Jet Images — Deep Learning Edition (Boosted W vs QCD)

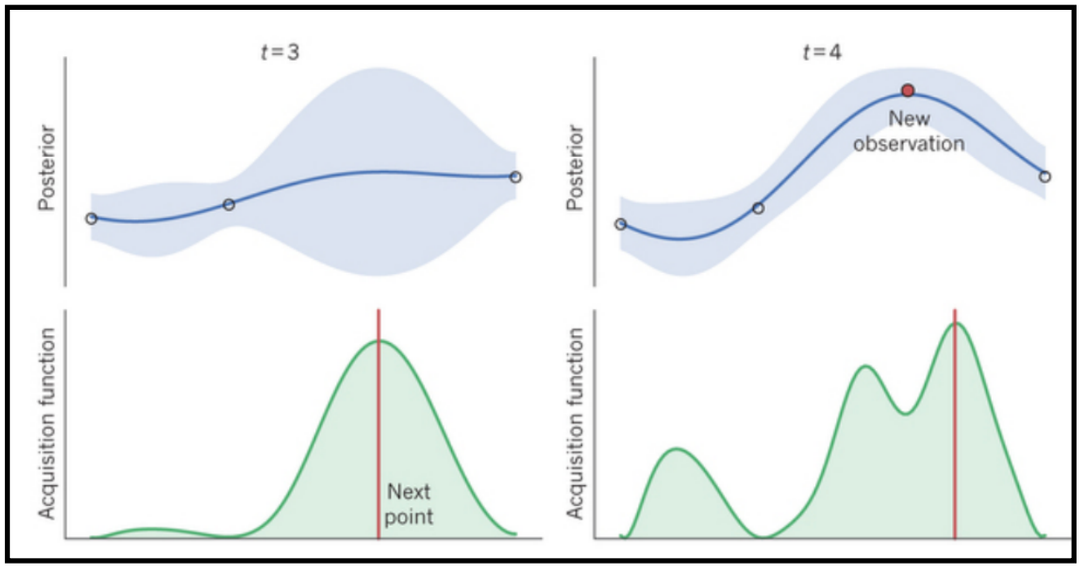


- Jet Constituents for Deep Neural Network Based Top Quark Tagging

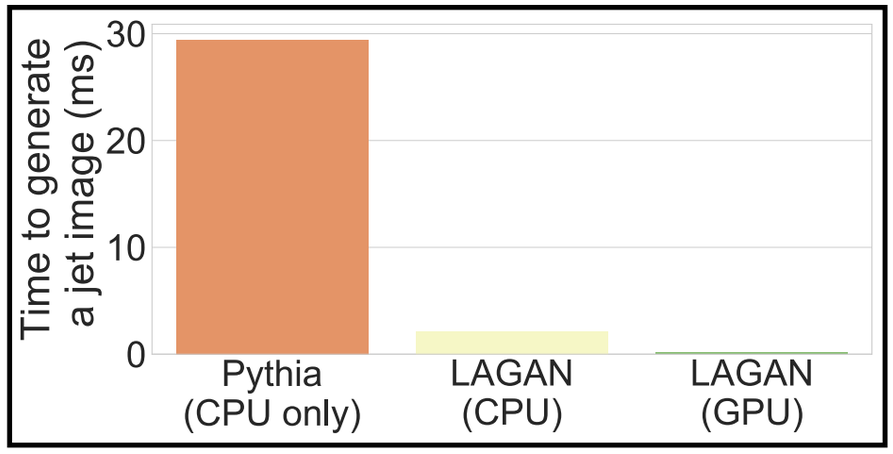


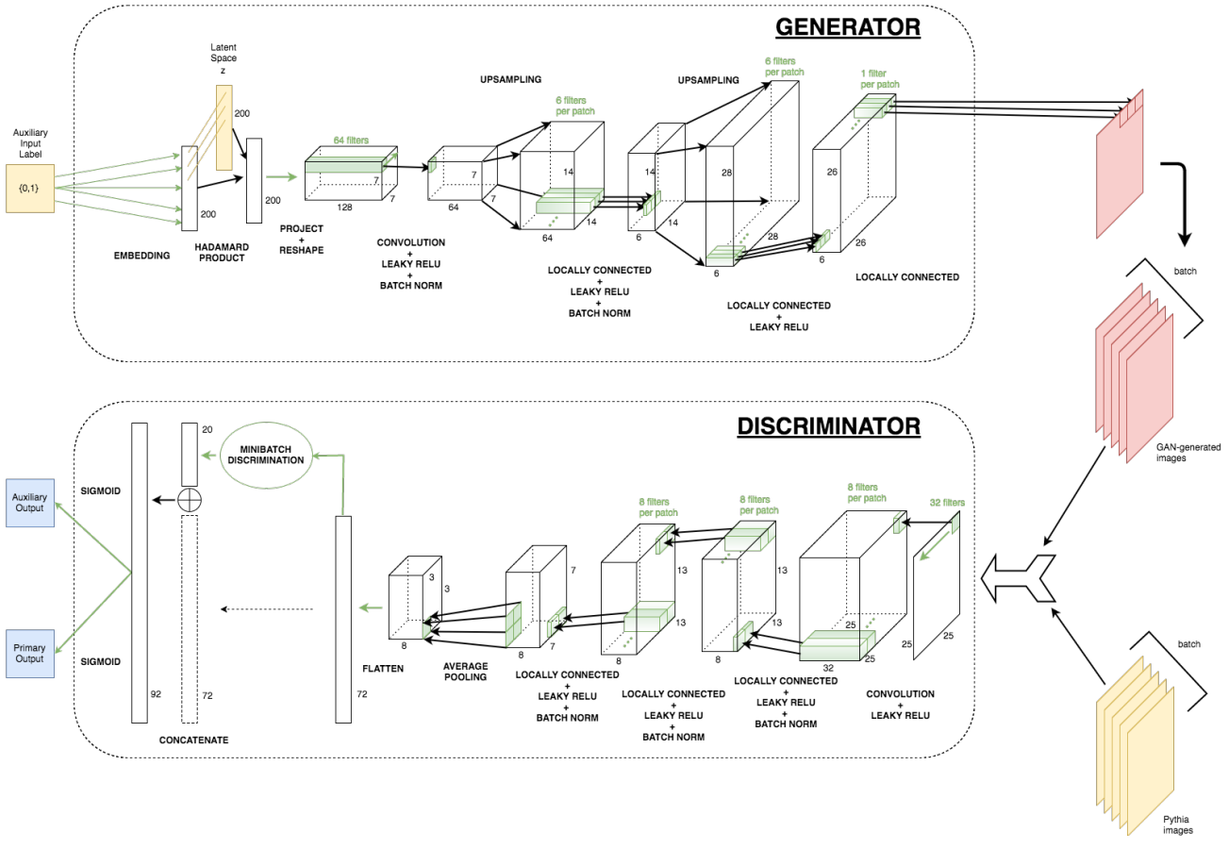
Generation

- Bayesian Optimization for Event Generation Tuning



- GANs (Generative Adversarial Networks) for Fast Simulation:





- Other efforts with Variational Auto-Encoders
- Future attempts may include Autoregressive Models

Yale

Computing

From Filippo Tilaro

- **Online monitoring**
  - Continuous service to analyse the system status and inform operators in case of fault detection
- **Fault diagnosis**
  - "Forensics" analysis of system faults that have already happened in the past. In some cases root-cause analysis
- **Engineering design**
  - Analysis of historical data to draw conclusions about system behaviours which could be helpful to improve / optimize the system under analysis

CERN openlab workshop 5

From Virginia Azzolini

WE NEED:
- operate effectively with minimal human guidance
- anticipate important events
- adapt behavior in response to changes in data content, user needs, or available resources

WE HOPE:
Developed machine learning and analytics-based solutions will:
- improve the accuracy of data quality
- establish new data quality rules to sharpen data error detection and correction
- increase the speed at which this is achieved

8

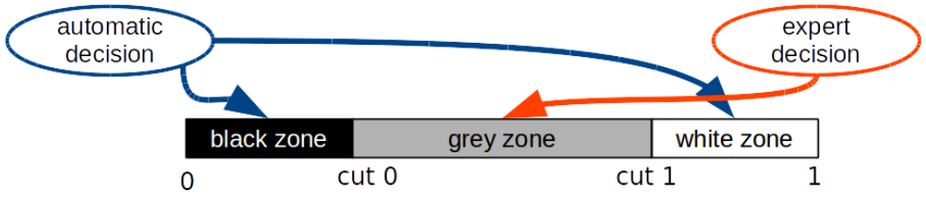
From Andrey Ustyuzhanin

Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree.

V. V. Gligorov[1] and M. Williams[2,a]

Reference Performance

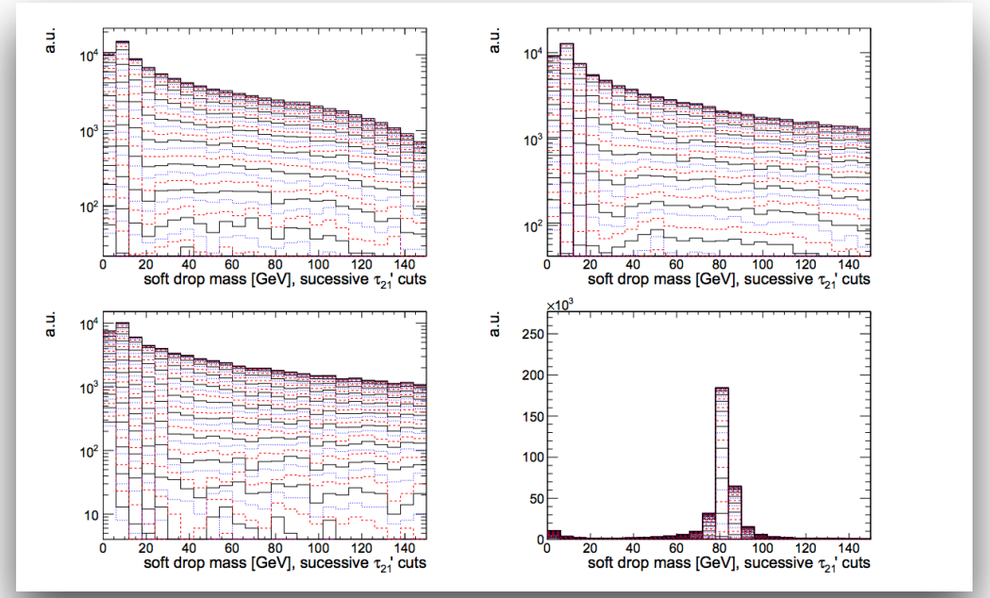
- Random Forest Classifier
- classification problem, 3 possible labels for each lumisection:
  - almost surely good ('white zone');
  - almost surely bad ('black zone');
  - ambiguous ('grey zone').



Decorrelating Discriminants

- Physics-driven method - DDT (designing decorrelated taggers)
- Can we use an adversary?

Dolan et al. arXiv:1603.00027

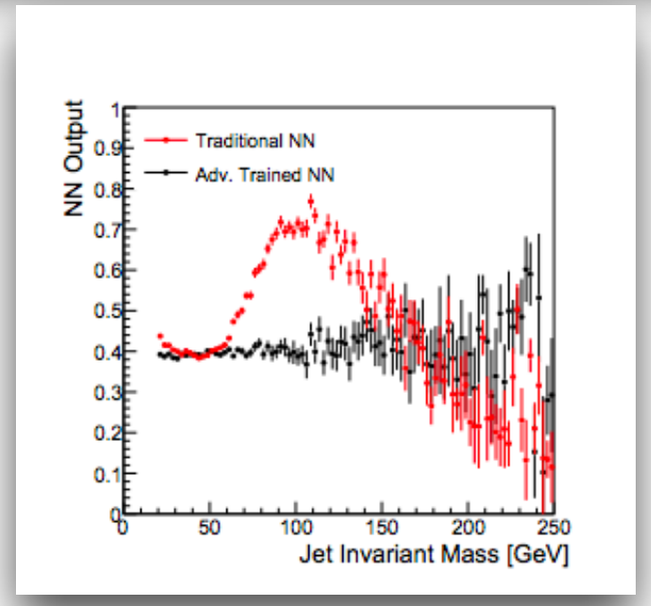


Learning to Pivot with Adversarial Networks

Gilles Louppe,[1] Michael Kagan,[2] and Kyle Cranmer[1]
[1] *New York University*
[2] *SLAC National Accelerator Laboratory*

Many inference problems involve data generation processes that are not uniquely specified or are uncertain in some way. In a scientific context, the presence of several plausible data generation processes is often associated to the presence of systematic uncertainties. Robust inference is possible if it is based on a pivot – a quantity whose distribution is invariant to the unknown value of the (categorical or continuous) nuisance parameters that parametrizes this family of generation processes. In this work, we introduce a flexible training procedure based on adversarial networks for enforcing the pivotal property on a predictive model. We derive theoretical results showing that the proposed algorithm tends towards a minimax solution corresponding to a predictive model that is both optimal and independent of the nuisance parameters (if that models exists) or for which one can tune the trade-off between power and robustness. Finally, we demonstrate the effectiveness of this approach with a toy example and an example from particle physics.

Louppe et al. arXiv:1611.01046



Shimmin et al. arXiv:1703.03507

What's next?

Continue improving in areas such as:

- Tracking
- Calorimetry
- Anomaly Detection
- so much more!