Artificial Intelligence – Tools and Applications in Particle Physics

Pai-hsien Jennifer Hsu
National Tsing Hua University

Outline

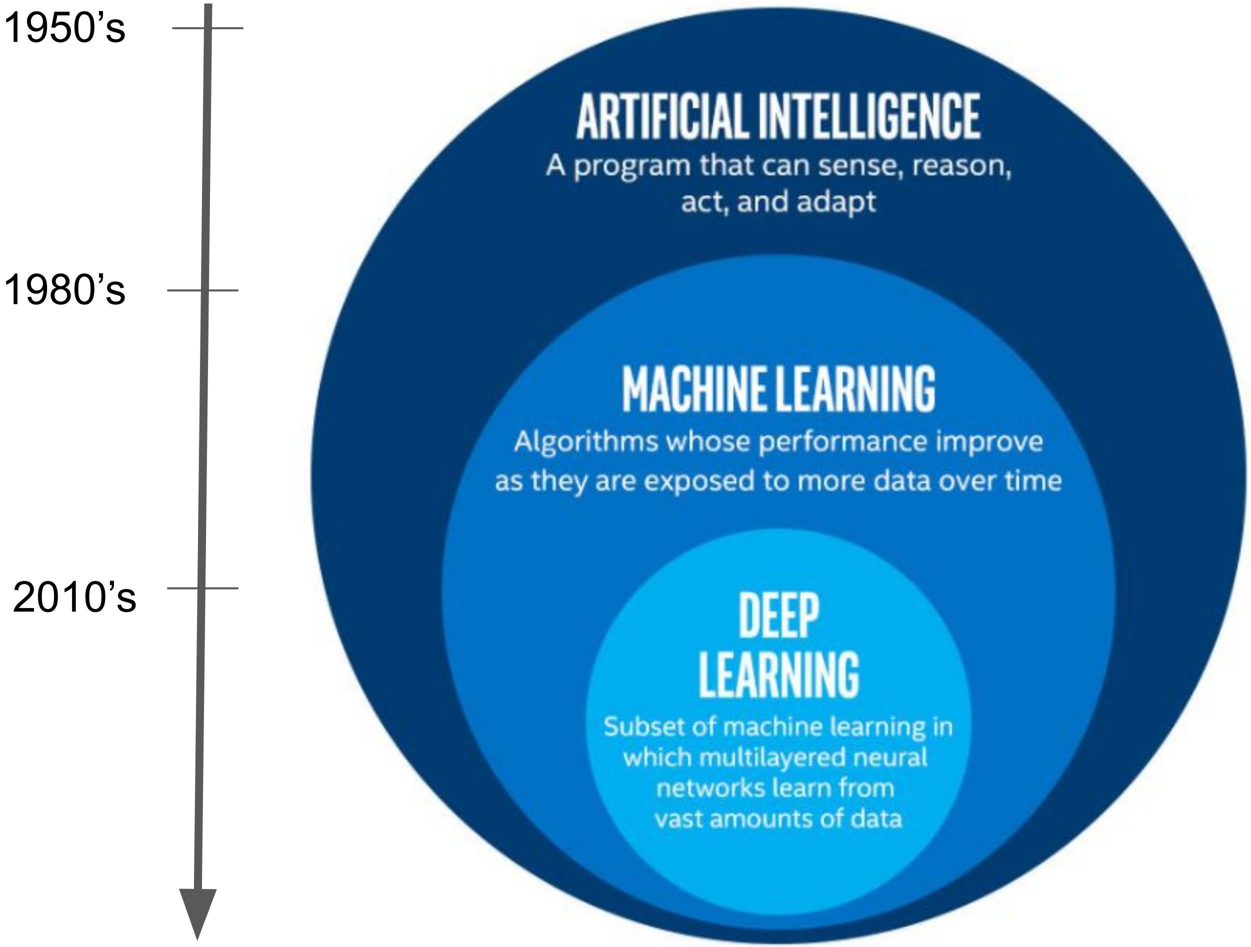
- What is Artificial Intelligence (AI)?
- What can AI do?
- What does Particle Physics need?
- What can AI do for Particle Physics?

AI?

Machine Learning?

What is AI?

Deep Learning?



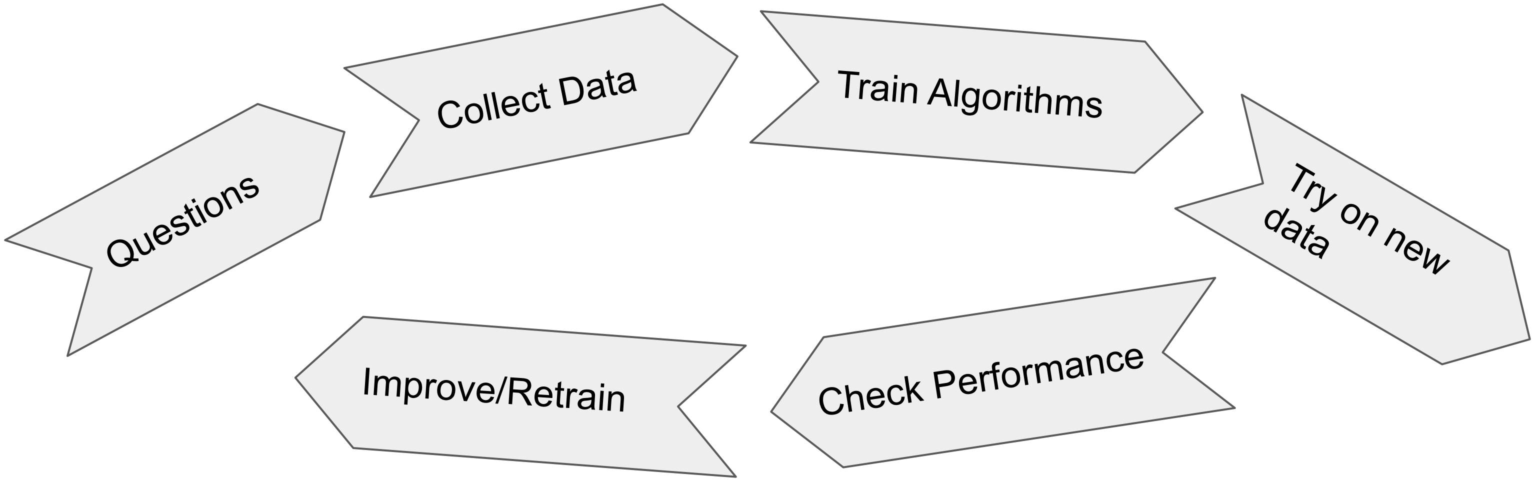
techplusmedia

Artificial Intelligence (AI)

- The science of training machines to perform human tasks
  - Understand human speech
  - Pattern/image recognition
  - Play strategic games (e.g. GO, chess)
  - Drive cars
- An ‘old’ idea since 1950’s

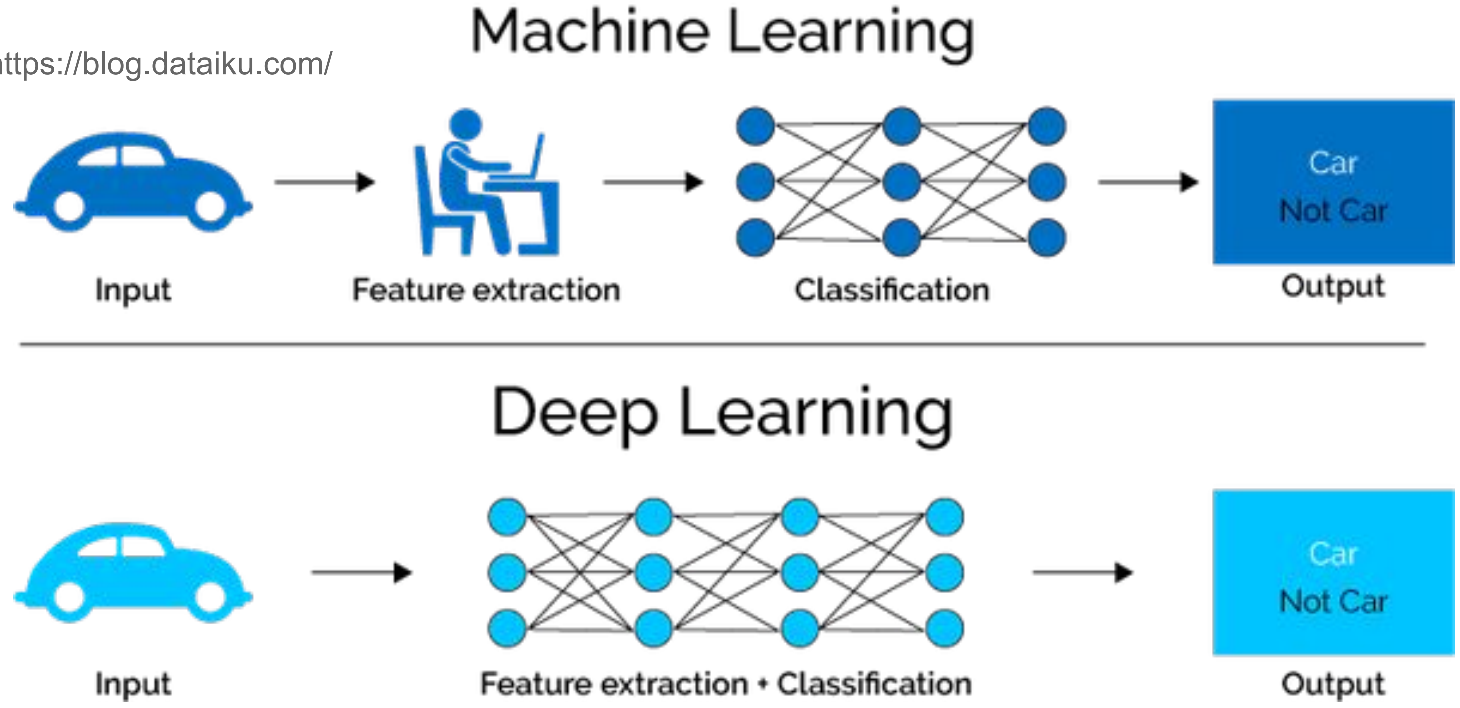
Machine Learning

- A subfield of AI which trains machine how to learn from data *without being explicitly programmed*



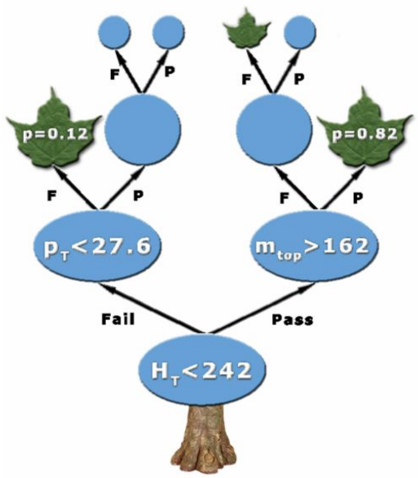
Classical Machine Learning vs Deep Learning

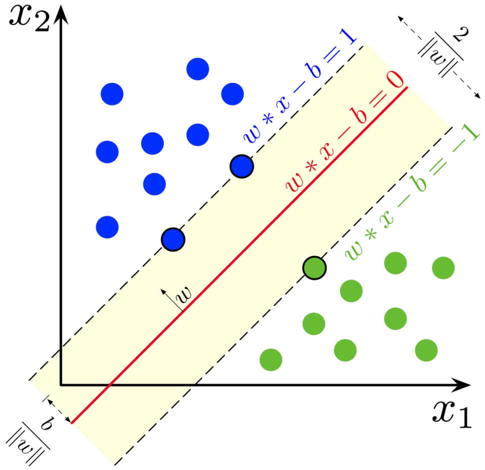
Source: https://blog.dataiku.com/



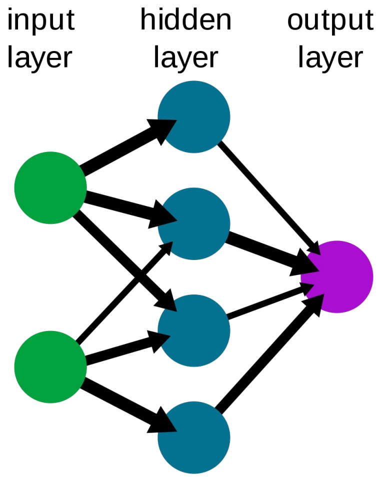
Classical Machine Learning vs Deep Learning

- ‘Classical’ machine learning typically refers to ‘simple’ models like decision trees, support vector machines, etc.
- Deep learning involves the use of artificial neural networks.





A simple neural network



Neural networks

- Different architectures and algorithms for various purposes.
  - Deep neural networks (DNN): versatile networks with many hidden layers (and neurons).

source

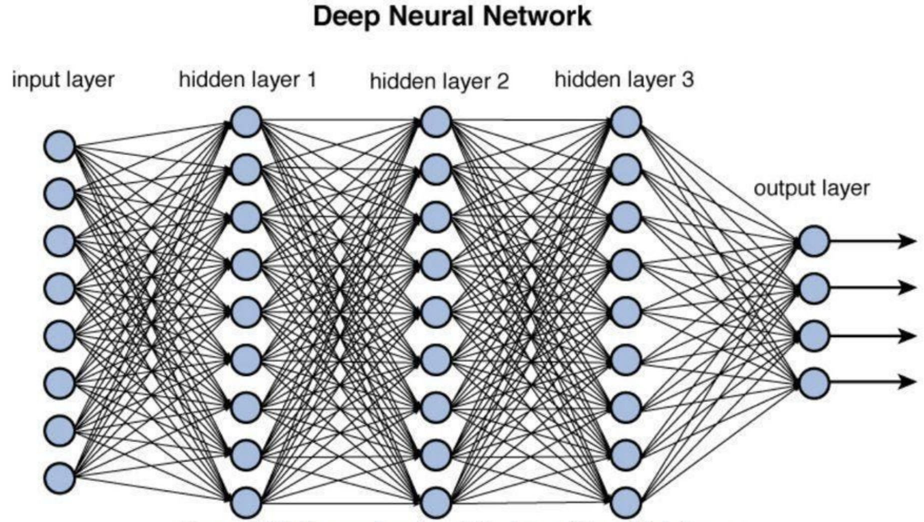
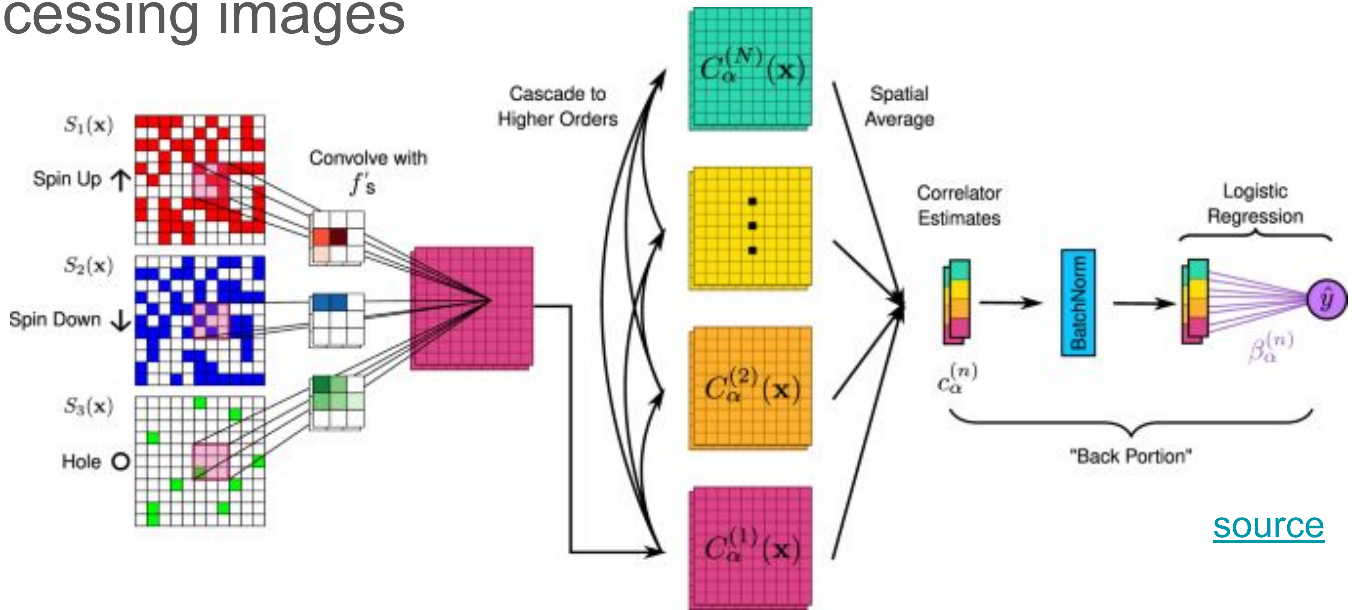


Figure 12.2 Deep network architecture with multiple layers.

Neural networks

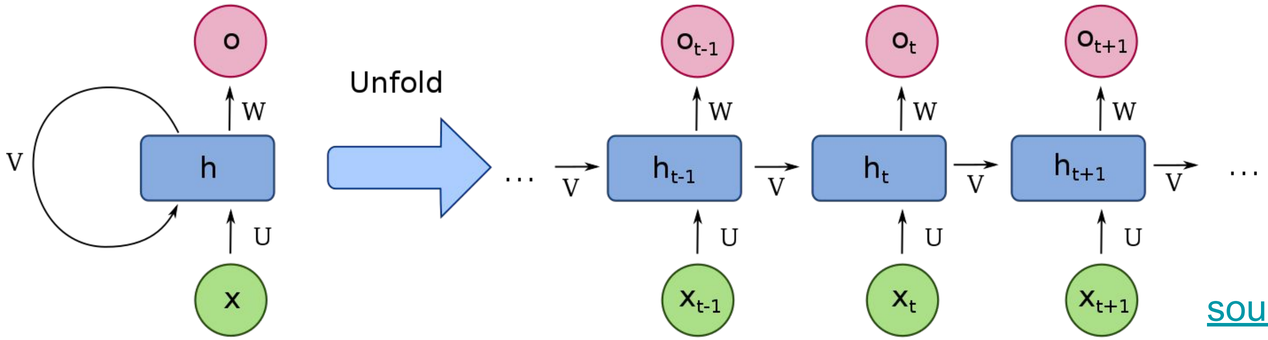
- Different architectures and algorithms for various purposes.
    - Convolutional neural networks (CNN): Powerful tool for processing images



source

Neural networks

- Different architectures and algorithms for various purposes.
    - Recurrent neural networks (RNN): 'backward' information flow; suitable for analyzing sequential data (like languages)



source

What can AI do?

Typical Tasks of AI

- AI models are built for specific tasks. For example,

  – Classification: Make discrete predictions (classify samples as True/False, or different categories.)

  – Regression: Make continuous predictions (given a feature X, predict the value of another feature Y).

  – Clustering: Separate data into different groups according to their similarities.

Typical methods of AI

- The methods used to train the AI model can be categorized as

  – Supervised learning

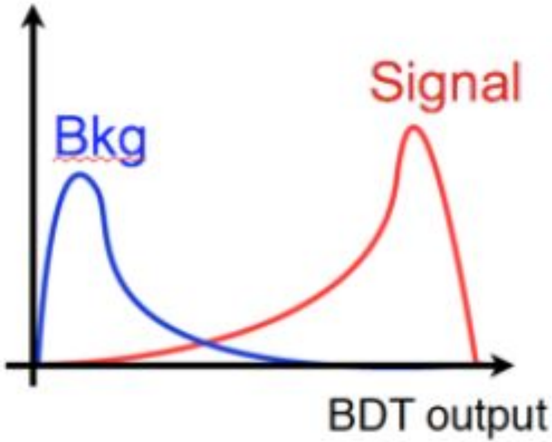
  – Unsupervised learning

Supervised Learning

- The right answer is given in the training data ('labeled')

  E.g. Train a boosted decision tree (BDT) with data labeled as signal and background

  Classification: to make *discrete* predictions (True/False, Signal/Background, Type I/II/III Supernovae, etc.)

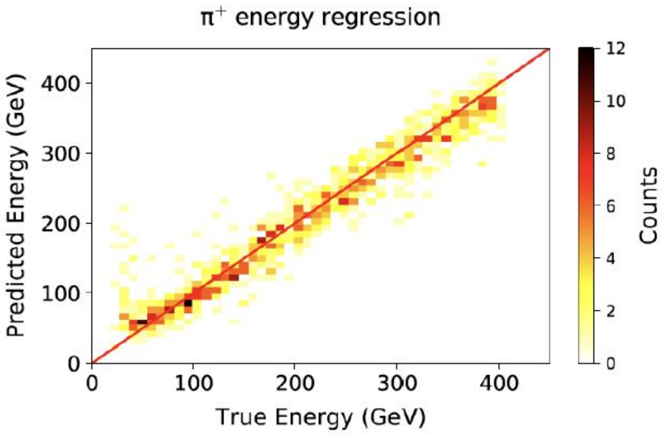


Supervised Learning

- The right answer is given in the training data (‘labeled’)

  E.g. Given the performance of the calorimeter, what is the true energy of a particle corresponding to a certain measured value.

  Regression: to make *continuous* estimation



Unsupervised Learning

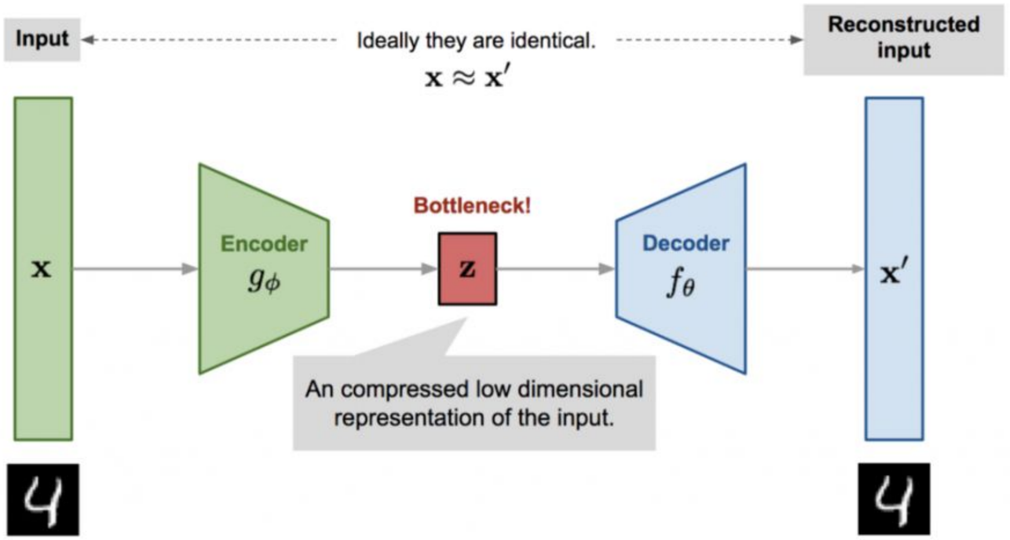
- The right answer is **not** given in the training data (‘unlabeled’)

  E.g. Given the observation, divide the stars into different groups

  Clustering: The groups are not known beforehand. Try to find the underlying structure of data.

Unsupervised Learning

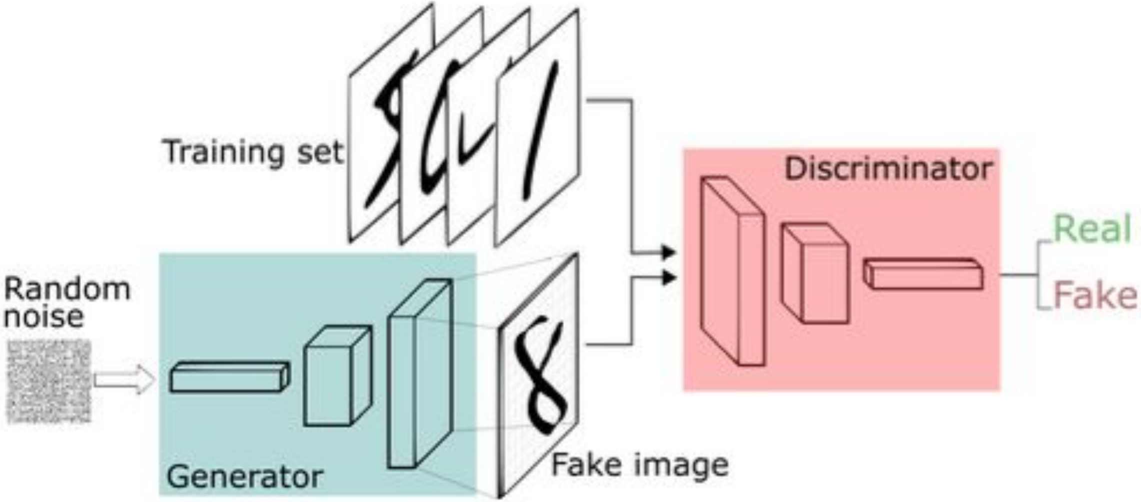
- Generative models: Generate (new/fake) data that look like input/training data. E.g. Autoencoders, Generative Adversarial Network (GAN)



Autoencoder

Unsupervised Learning

- Generative models: Generate (new/fake) data that look like input/training data. E.g. Autoencoders, Generative Adversarial Network (GAN)



GAN

Machine Learning

- Key ideas: identify a problem you want to solve, and learn from data
  - **Domain knowledge**: Understand what the problem is, and what information are needed to solve the problem
  - To apply AI/Machine learning to particle physics, you should first be a ‘domain expert’ of particle physics!

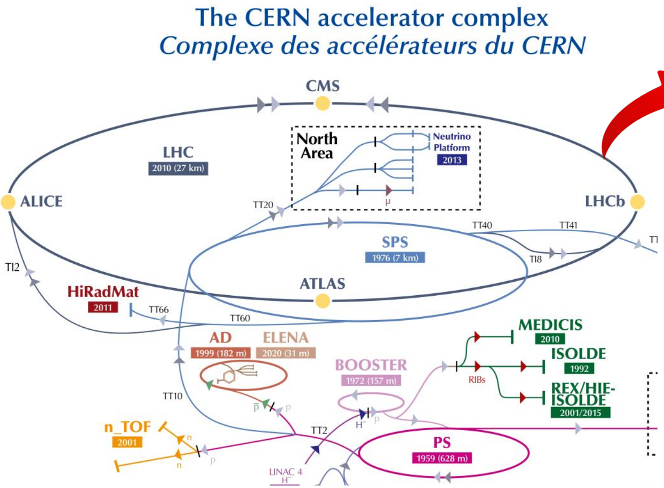
What does Particle Physics need?

Particle Physics

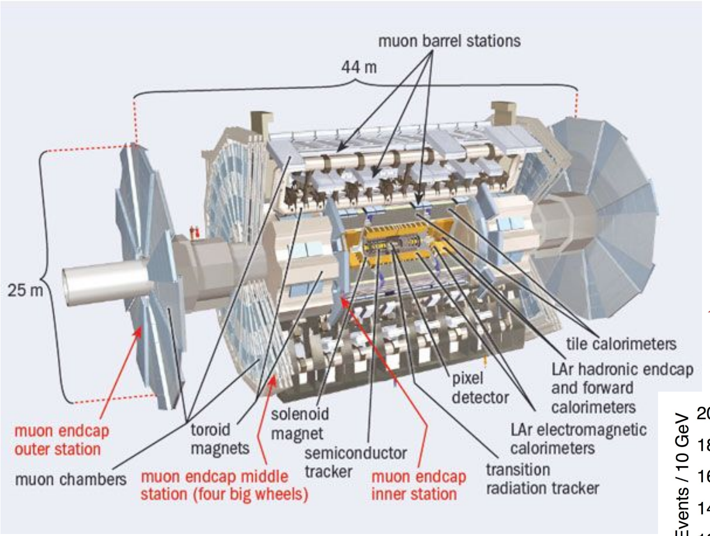
- Particle physics is a broad topic. Focus on experimental particle physics, in particular collider physics, for this lecture.

  Example: Large Hadron Collider (LHC) Experiments

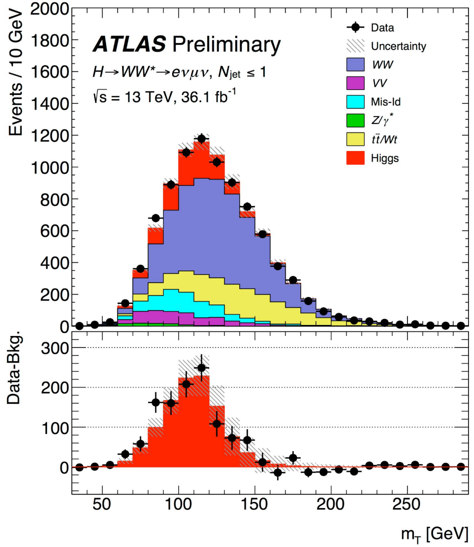
Example: LHC



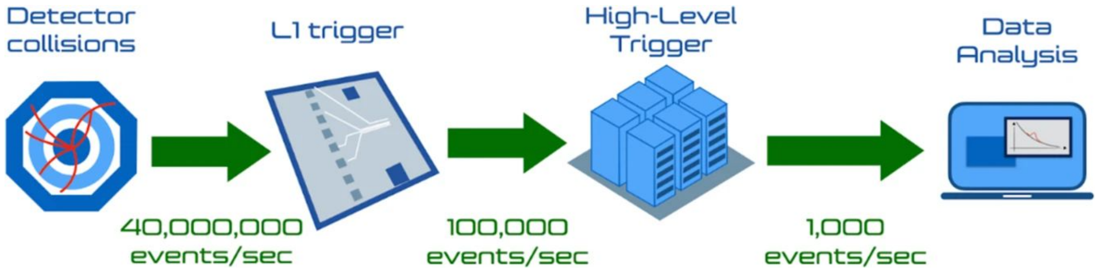
The CERN accelerator complex
Complexe des accélérateurs du CERN
CMS
LHC
2010 (27 km)
North Area
Neutrino Platform
2013
ALICE
LHCb
TT20
TT40
TT41
SPS
1976 (7 km)
TI8
TI2
HiRadMat
2011
TT66
ATLAS
TT60
AD
1999 (182 m)
ELENA
2020 (31 m)
MEDICIS
2010
ISOLDE
1992
BOOSTER
1972 (157 m)
RIBs
REX/HIE-ISOLDE
2001/2015
TT10
n_TOF
2001
TT2
PS
1959 (628 m)
LINAC 4



muon barrel stations
44 m
25 m
tile calorimeters
LAr hadronic endcap and forward calorimeters
pixel detector
LAr electromagnetic calorimeters
solenoid magnet
toroid magnets
muon endcap outer station
semiconductor tracker
transition radiation tracker
muon chambers
muon endcap middle station (four big wheels)
muon endcap inner station



ATLAS Preliminary
$H \to WW^* \to e\nu\mu\nu$, $N_{jet} \leq 1$
$\sqrt{s}$ = 13 TeV, 36.1 fb^{-1}
Data
Uncertainty
WW
VV
Mis-Id
Z/γ^*
$t\bar{t}/Wt$
Higgs
Events / 10 GeV
Data-Bkg.
m_T [GeV]



Detector collisions
L1 trigger
High-Level Trigger
Data Analysis
40,000,000 events/sec
100,000 events/sec
1,000 events/sec

Needs of Particle Physics

- Produce the data
- Acquire the data
- Clean the data
- Manage the data
- Reconstruct the data
- Simulate the data
- Analyze the data
- Find the theory behind the data

We can make use of AI in all these areas!

What can AI do for Particle Physics?

Applications of AI in Particle Physics

- Due to limited time only a few selected examples are presented. More information (live update) at HEPML-LivingReview

Production of data

- We use accelerators to speed up the particle beams, collide them at designated points, and measure the scattering particles.

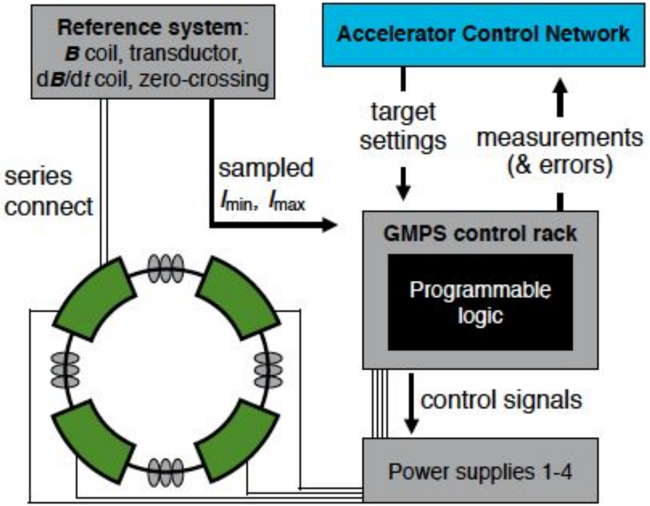
- We use AI/machine learning to
  – Predict and avoid equipment failures
  – Control the beams and optimize the quality of the beams

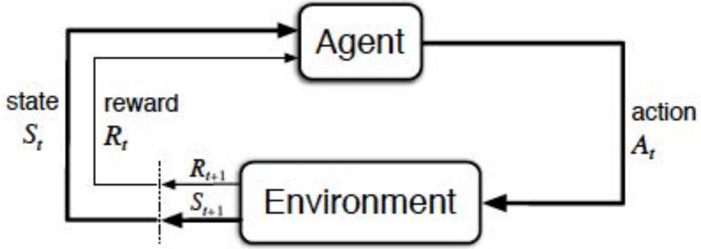
Production of data

arXiv:2011.07371

- R&D going on: how to run accelerators in a more efficient, more reliable, and possibly even autonomous way?

- E.g. Reinforcement learning to regulate power supply at Fermilab





Acquisition of data

- Data production rate at collider experiments are so high (e.g. ~40 TB/s at the LHC) that one cannot save all the data recorded.

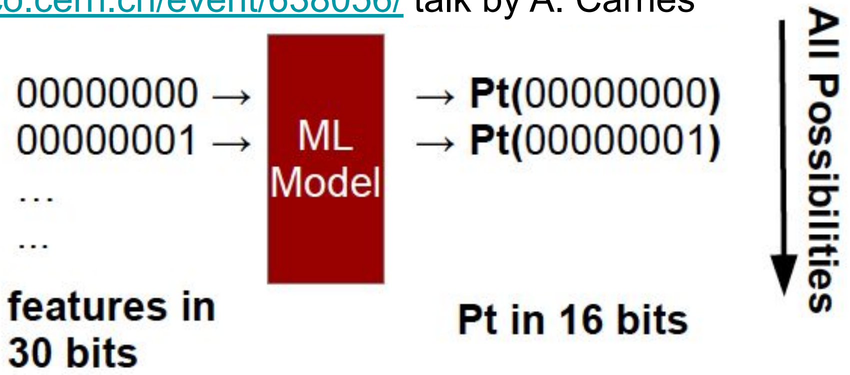
- Rely on trigger systems to filter events to a manageable amount for storage. Challenge: How to make sure we've kept the 'interesting' events?
  – AI can help!

Acquisition of data

- Example: Muon Trigger at CMS
  – Need to predict momentum of muons correctly while keep the event rate as low as possible.
  – A regression tasks with many features

https://indico.cern.ch/event/638056/ talk by A. Carnes

– ML Model = Boosted Decision Tree
– Used in 2016/2017 data taking.



Data Quality Monitoring

- Need to use high quality data, i.e. data recorded by a ‘healthy’ detector, for subsequent physics and performance analyses.

- Important to spot detector issues in real time to avoid taking low-quality data → Data Quality Monitoring is crucial.

- Complicated sub-detectors/sub-systems involved. Can AI help?

source

Data Quality Monitoring

https://doi.org/10.1051/epjconf/201921406008

- Example: Anomaly detection with autoencoder for data quality monitoring in the CMS experiment

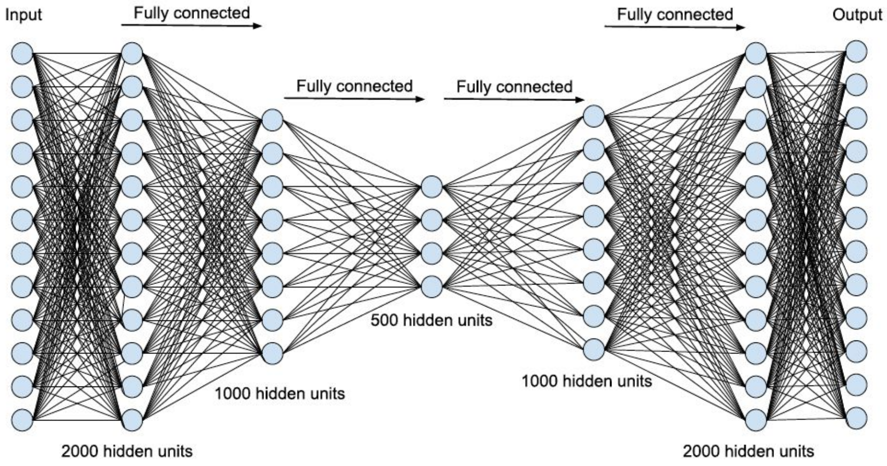
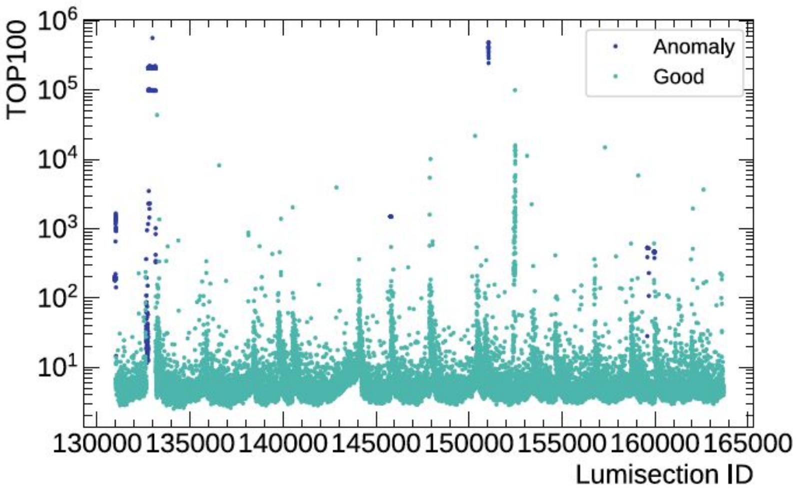
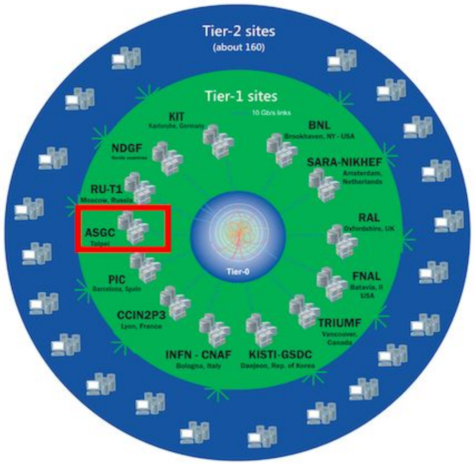


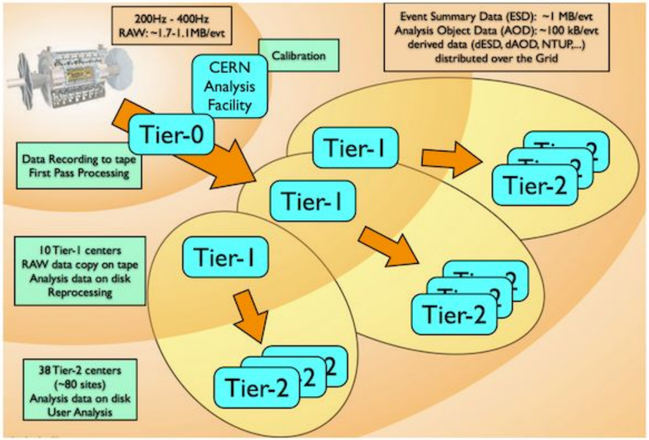
Figure 1. Proposed autoencoder architecture.



Data Management

- LHC experiments rely on the worldwide computing grid for data storage, distribution, access, and processing.
  – AS Grid computing center is one of the Tier-1 grid sites





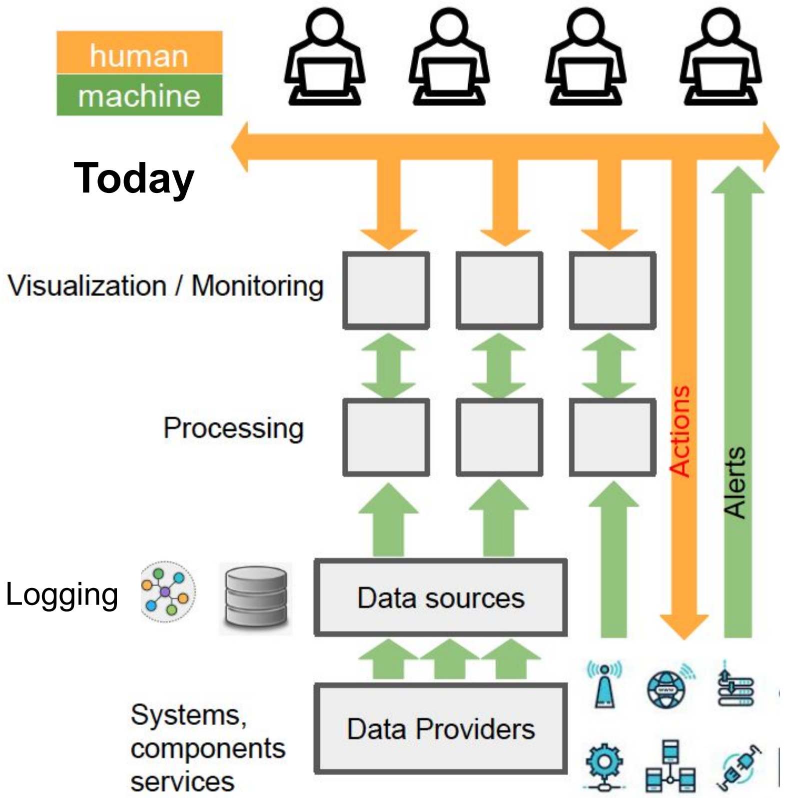
TW Grid

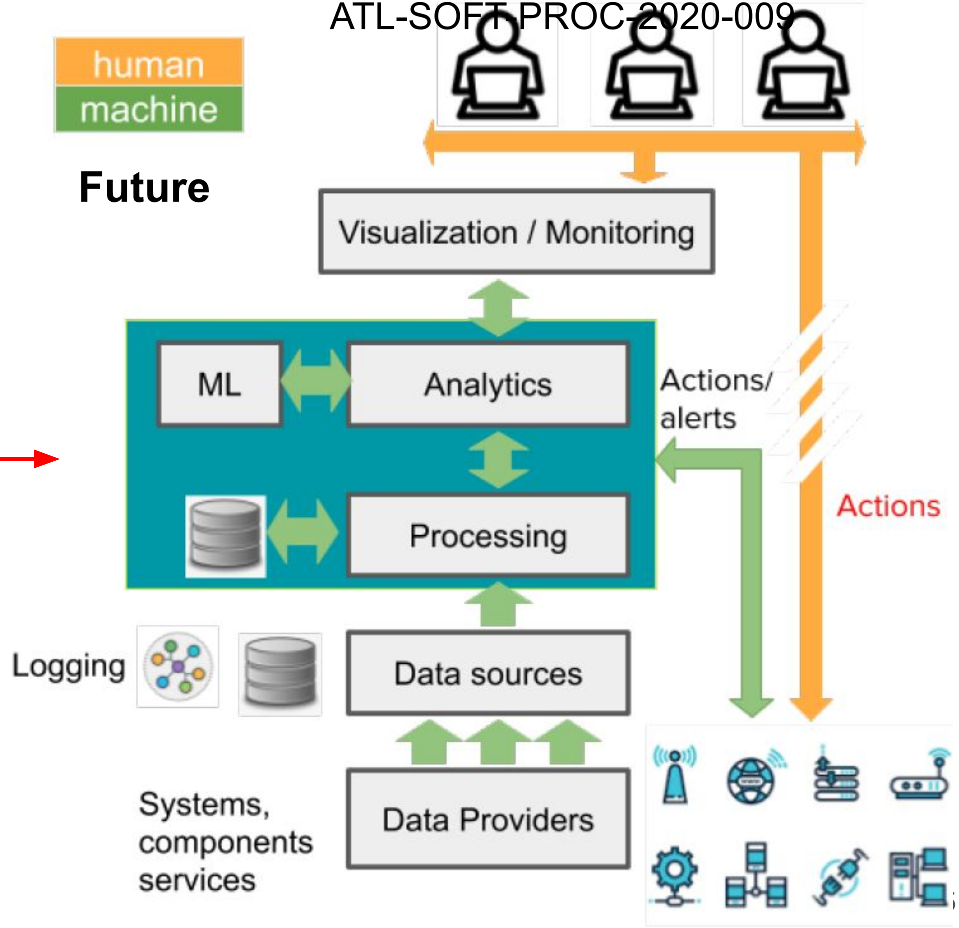
Data Management

- Grid maintenance is persen-power demanding. → Use AI to improve efficiency and reduce human effort.

- Example: The Operational Intelligence project (a joint effort across the LHC worldwide grid communities).

Data Management

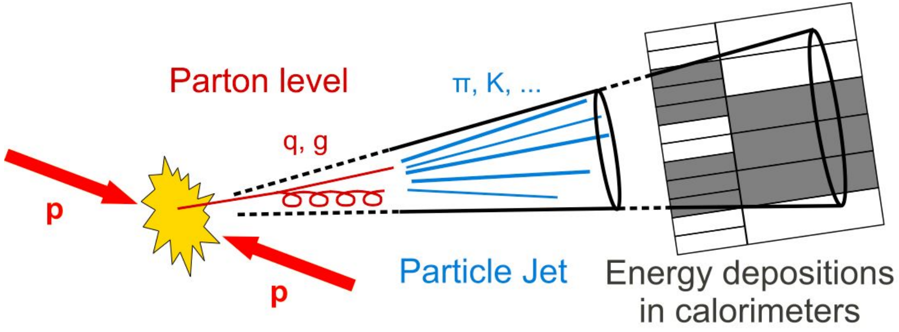




Reconstruction of Data

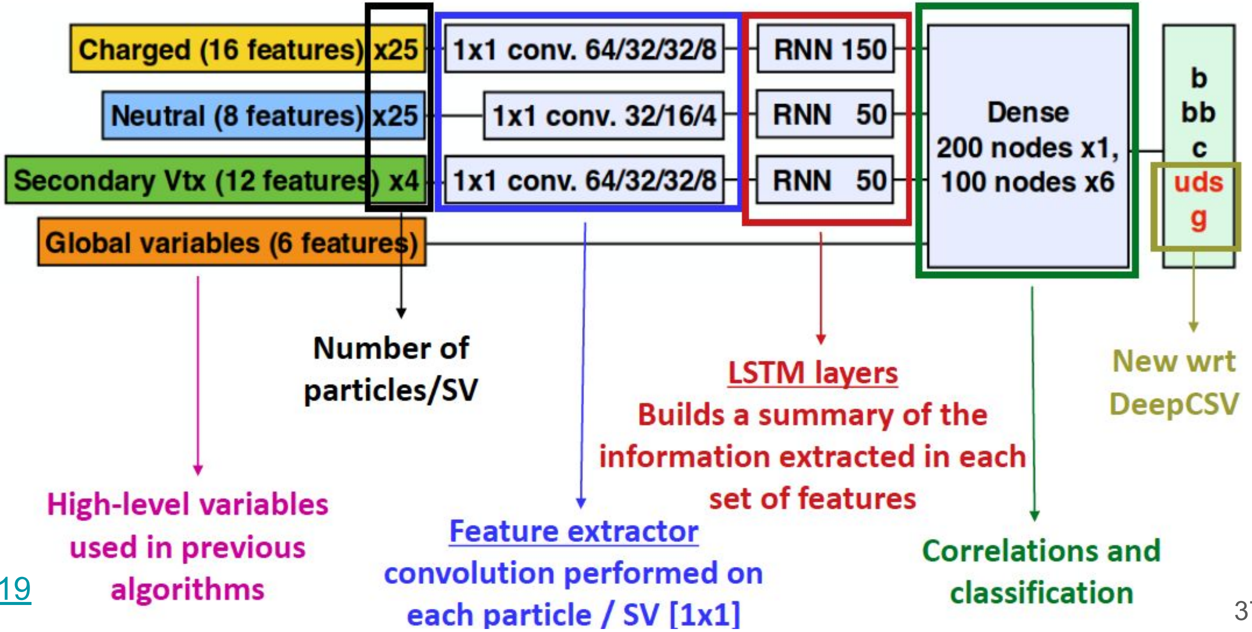
- How to identify the particles created from collision using information recorded by the detector?

- E.g. Jet = a cone of reconstructed particles. What's its origin? Quark (which flavor)? Gluon? Hadron? → **Flavor Tagging**



Reconstruction of Data

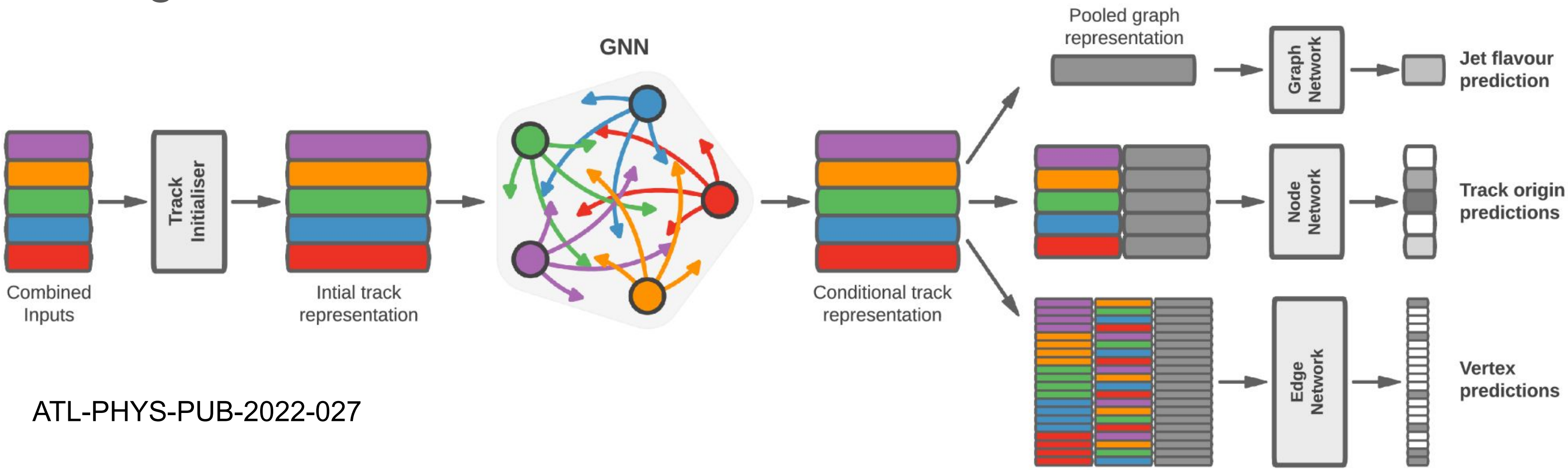
- Treat a jet as a sequence of constituent particles and use Deep neural networks, e.g. DeepFlavour in CMS



L. Gouskos, CEPC 2019

Reconstruction of Data

- Treat a jet as a graph and make use of Graph Neural Network. E.g. GN1 in ATLAS



ATL-PHYS-PUB-2022-027

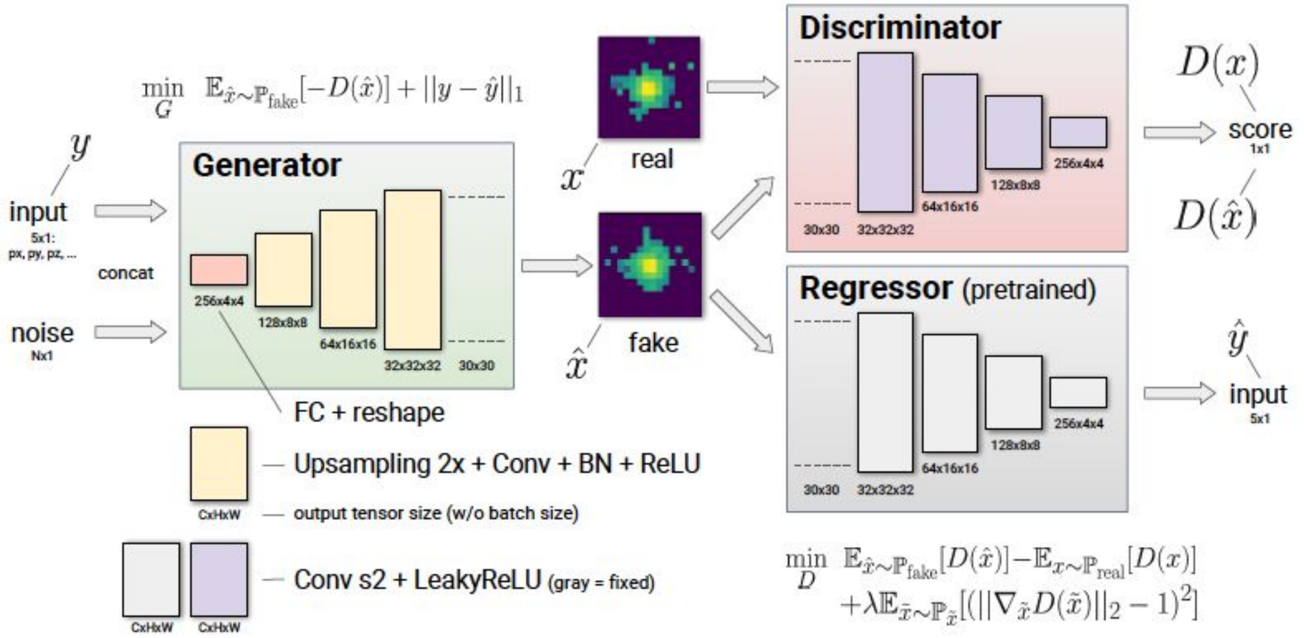
Simulation of Data

- We rely on simulated data to compare real collision data with known physics.
- Need to accurately describe the underlying physics and detector responses. → Computationally expensive.
- Idea: train a generative model, and use it as a statistical model embedded with high-level detector response.

Simulation of Data

- Example: GAN for LHCb fast simulation

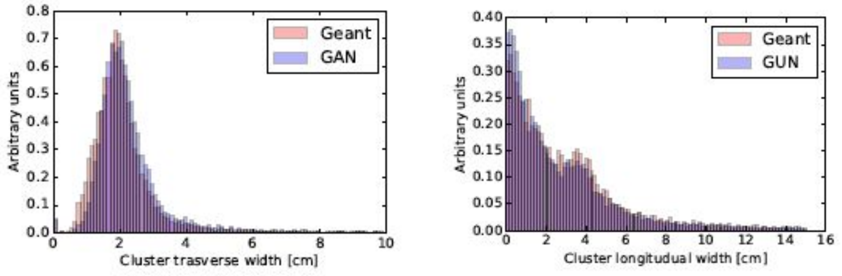
arXiv:2003.09762



Simulation of Data

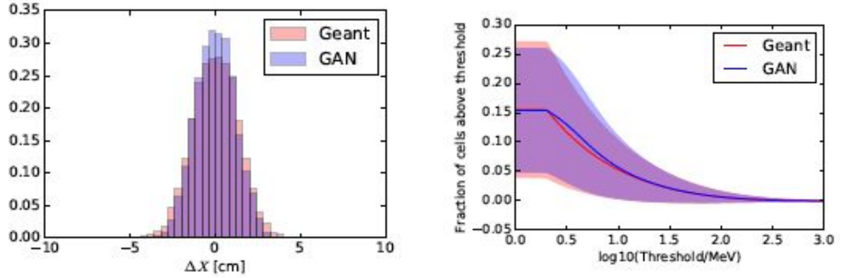
arXiv:2003.09762

- Example:
  GAN for LHCb fast simulation



(a) The transverse width of clusters.

(b) The longitudinal width of clusters.

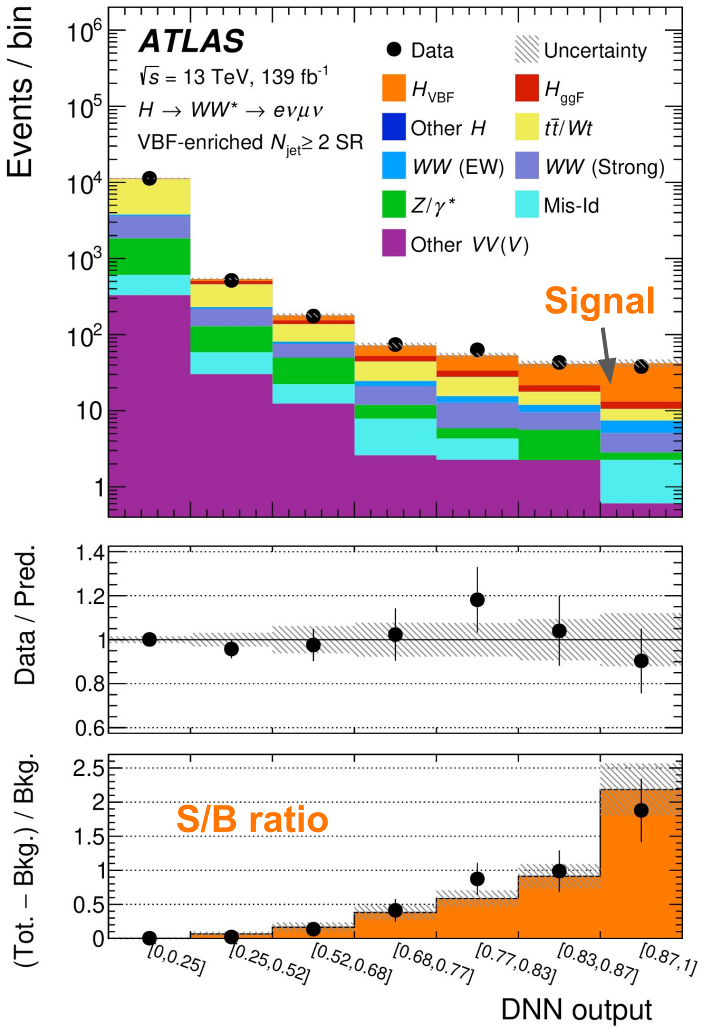


(c) ΔX between cluster centre of mass and the true particle coordinate.

(d) The sparsity in the 30x30 cells matrix containing clusters.

Analysis of Data

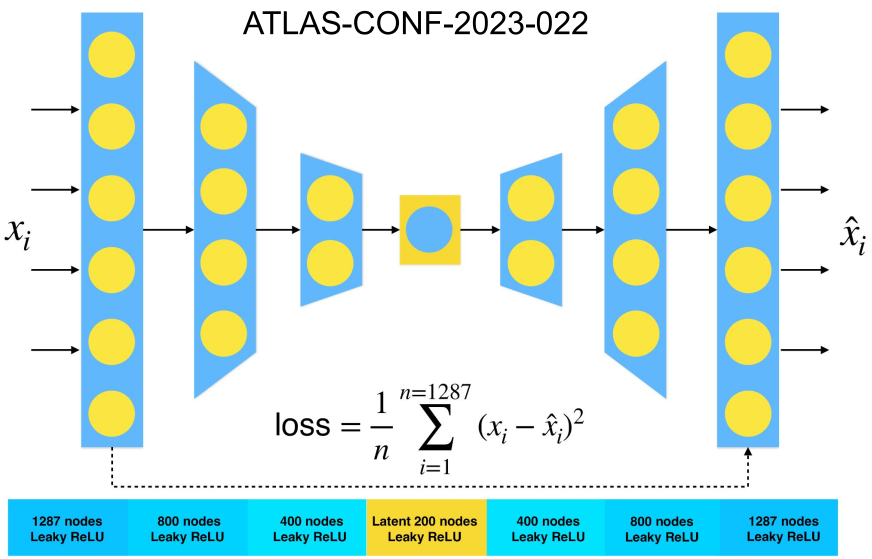
- AI is extremely effective in classification (separation of signal and background events) → extensive use of DNN in analysis.

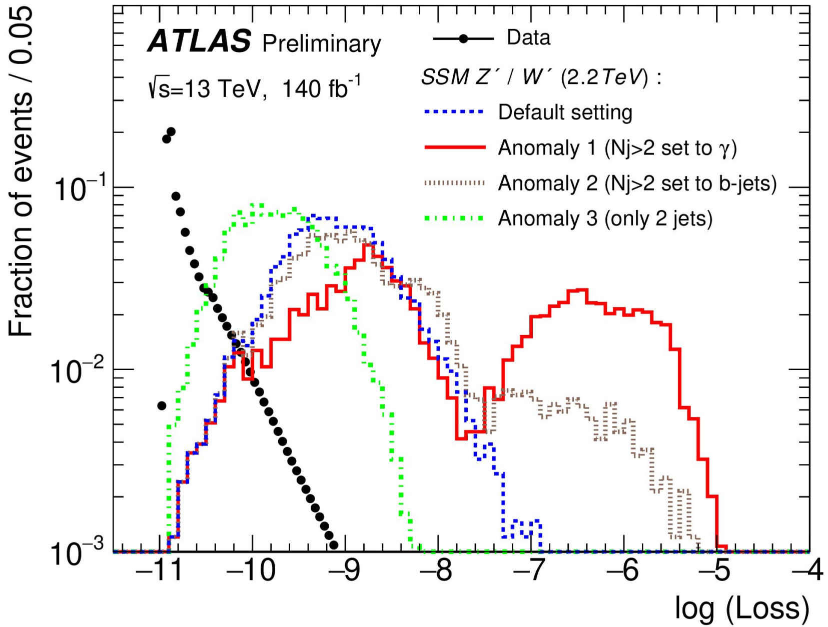


arXiv:2207.00338

Analysis of Data

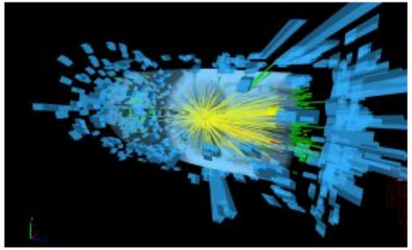
- Model-independent search: Anomaly detection





Theory behind the Data

- The main part of data analysis in HEP is to do hypothesis test: does it look more like the SM/background (‘hypothesis 0’) or the signal/new physics (‘hypothesis 1’)?
  – Need to evaluate the likelihood

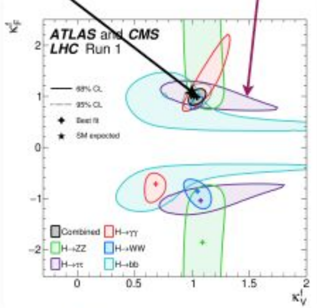


High-dimensional event data x

Likelihood function $p(x|\theta)$

Maximum-likelihood estimator

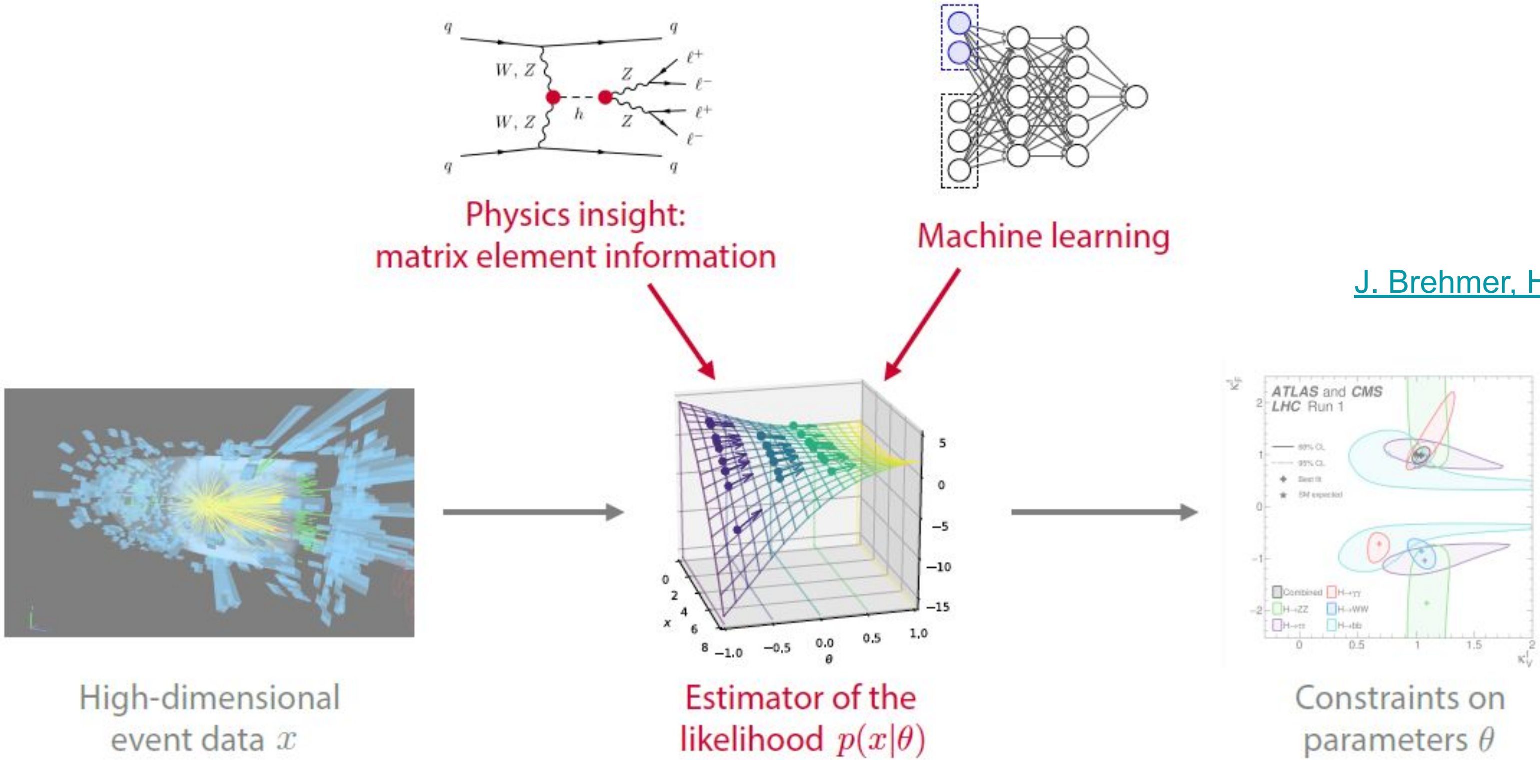
Confidence limits based on likelihood ratio tests



Constraints on parameters θ

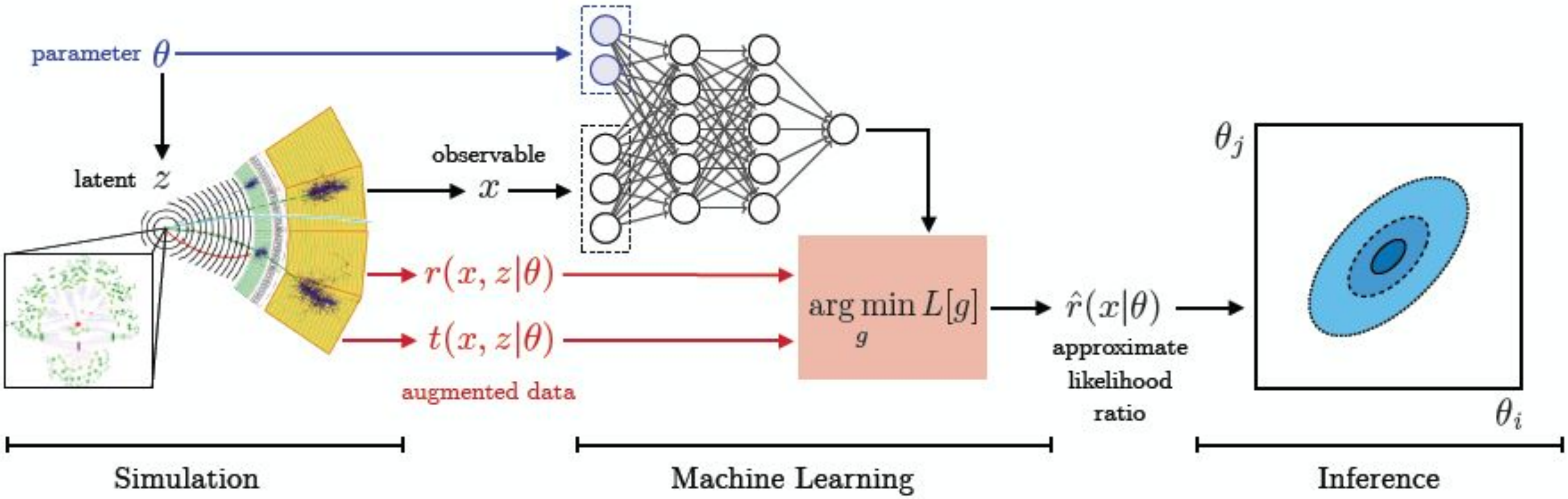
J. Brehmer, HEFT 2019

Theory behind the Data



J. Brehmer, HEFT 2019

Theory behind the Data



arXiv:1805.00013

Take-home message

- High energy experiments are full of big data, and opportunities for computing intensive innovations. → Rapid growth of AI applications in particle physics.

- More R&D ongoing: how to make use of symmetries (of the experimental setup and the underlying physics), how to keep systematics under control, how to better understand/explain results obtained by AI, etc. → Exciting AI+Physics era ahead!