ML in Particle Physics

Eilam Gross
with Jonathan Shlomi & Sanmay Ganguly

Refs:

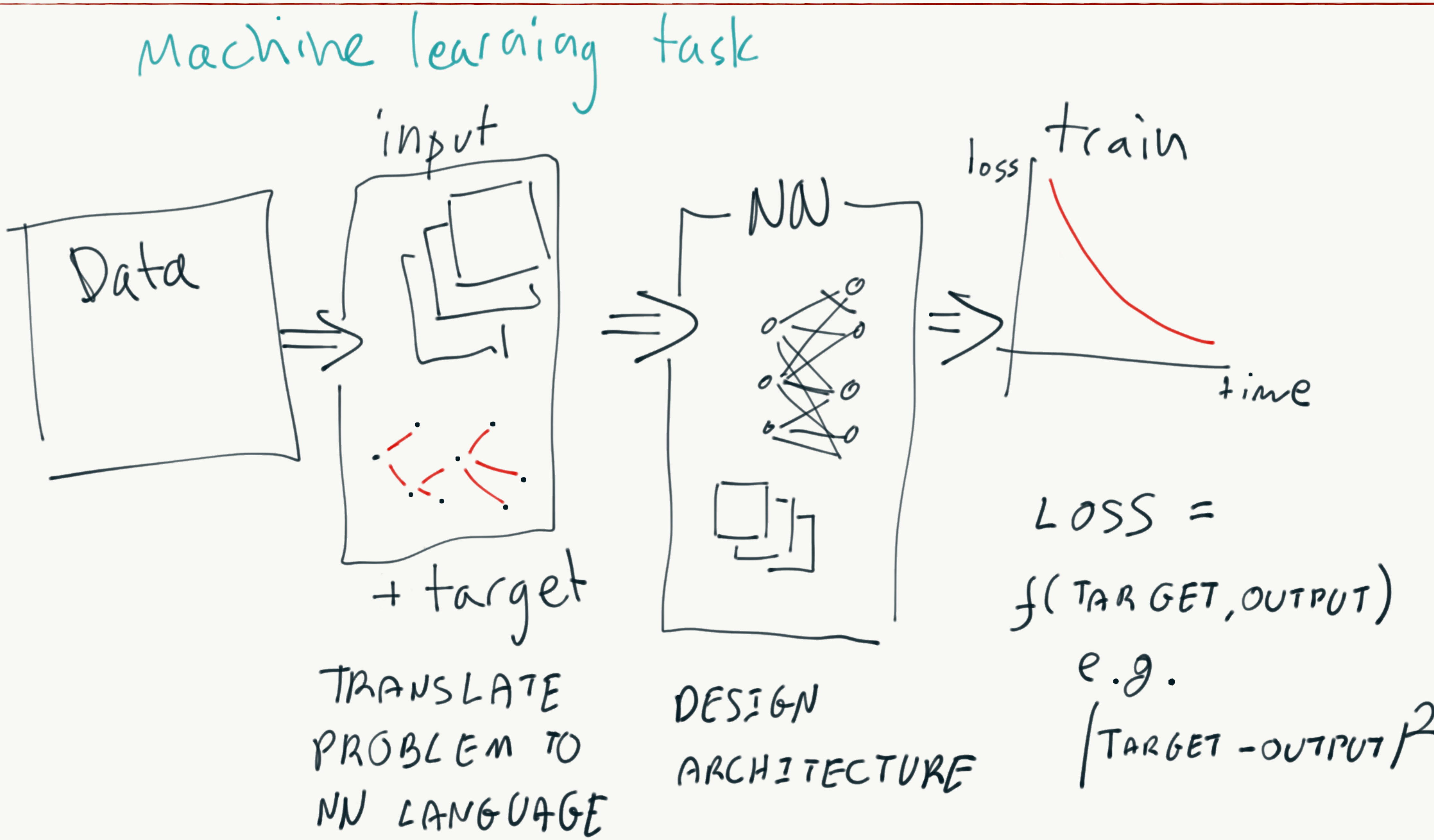
M. Pierini - private correspondence

Machine Learning in High Energy Physics Community White Paper,
https://arxiv.org/abs/1807.02876

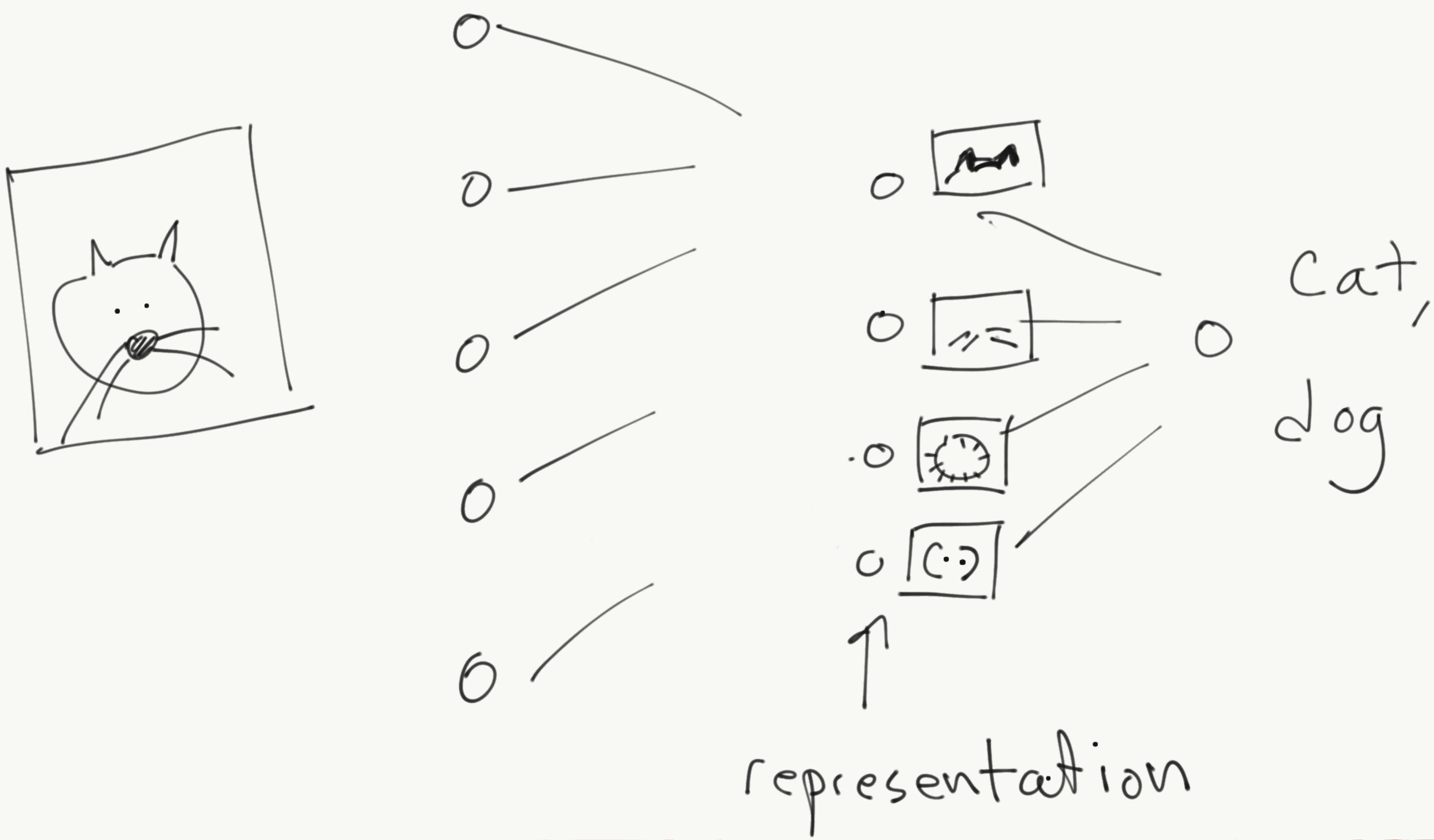
ML: How Does it Work?

Machine Learning (ML) is a branch of computer science that develops techniques to train algorithms to accomplish a task, without explicitly programming the algorithms to do so.

ML algorithm is provided with a loss function to minimize (quantifying how badly the task is accomplished) and a training dataset to learn how to accomplish the task at best.



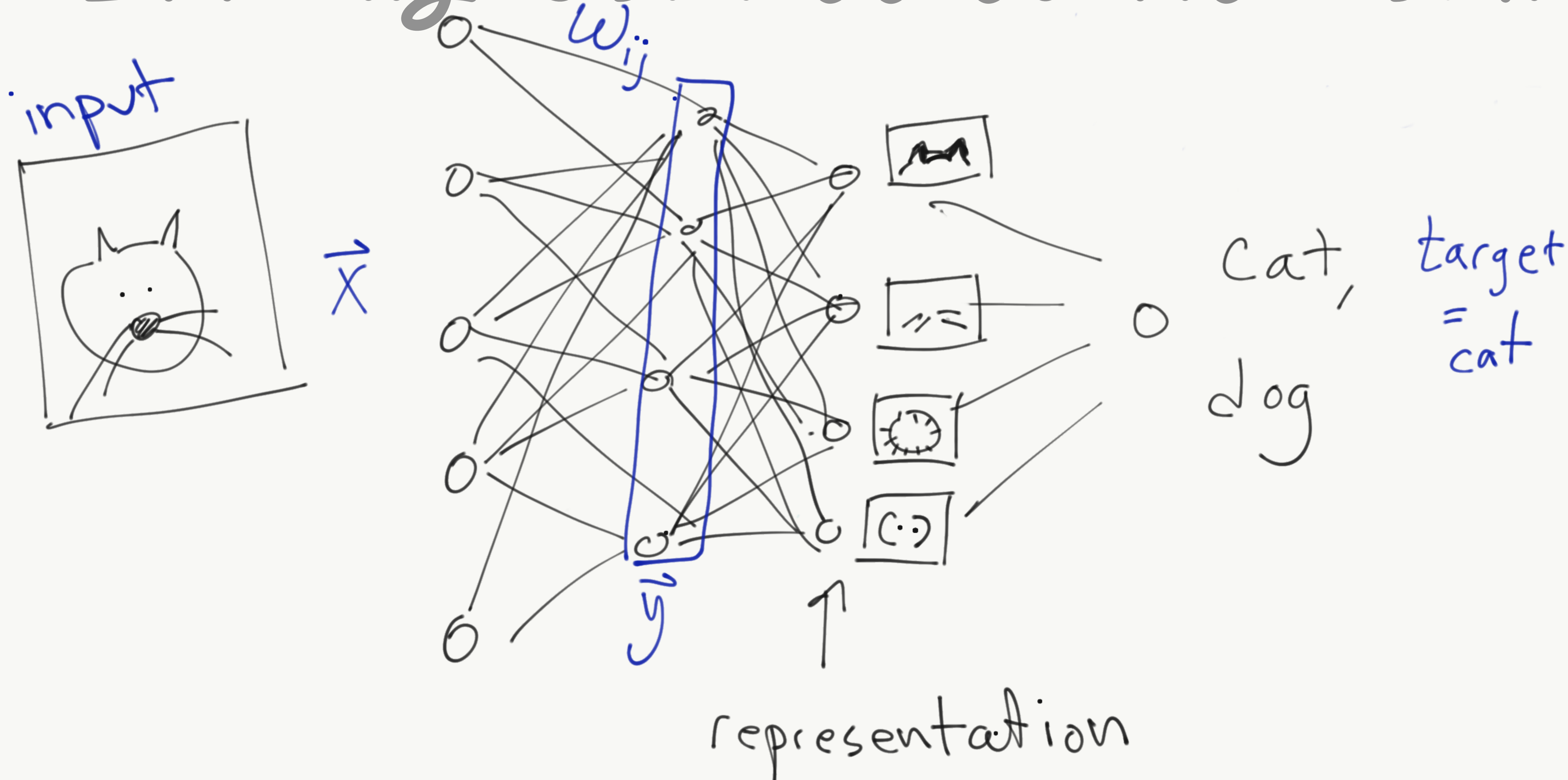
ML: How Does it Work?



The representation is like a language
the machine is creating to translate the input image
so the desired targeted output can be achieved
in an optimised way

The representation is implemented with an intelligent
human architecture

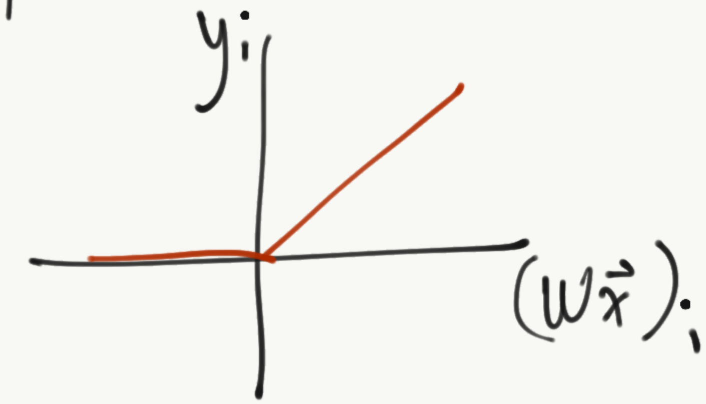
DL: Fully Connected Network



y is the input to the next layer

$$\vec{y} = g(W\vec{x})$$

$g =$



g = activation function

Loss = (target - input)^2

Use $\vec{\nabla} L = \frac{\partial L}{\partial w_{ij}}$

The training stage is most expensive.
The inference stage is fast and inexpensive since the weights were determined

Why Machine Learning

- ML methods are designed to exploit large/multi-dimensional DATA sets in order to reduce complexity and find new features in the DATA

- HEP goes into the realm of high luminosity, which means BIG DATA

Machine Learning and HEP

- BDT and NN used in HEP for quite some time
- Typically, variables relevant to the physics problem are selected and a machine learning model is trained for classification or regression using signal and background events (or instances).
    - Classification (predicting a label)
        - Particle ID
        - Heavy flavor tagging
        - top/W/H Tag
    - Regression (predicting a continuous quantity)
        - Use cluster shape information to improve calorimeter energy resolution
        - Pileup estimation

ML and HEP

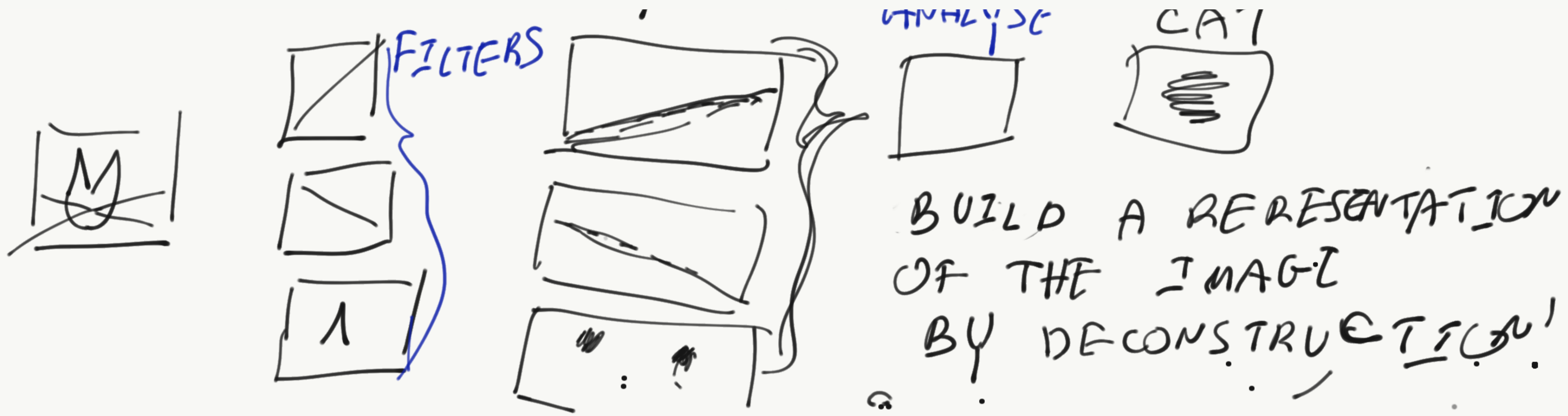
- But ML can do much more than that
- HEP is not only about classification or finding new Physics. It is also about Quality of Data and Detectors, Simulations and triggering, where ML is starting to take bigger roles.
- Increase computer power and sophistication in algorithms brought the Deep Learning revolution of the last decade, which is largely affecting HEP

Deep Learning in a NutShell

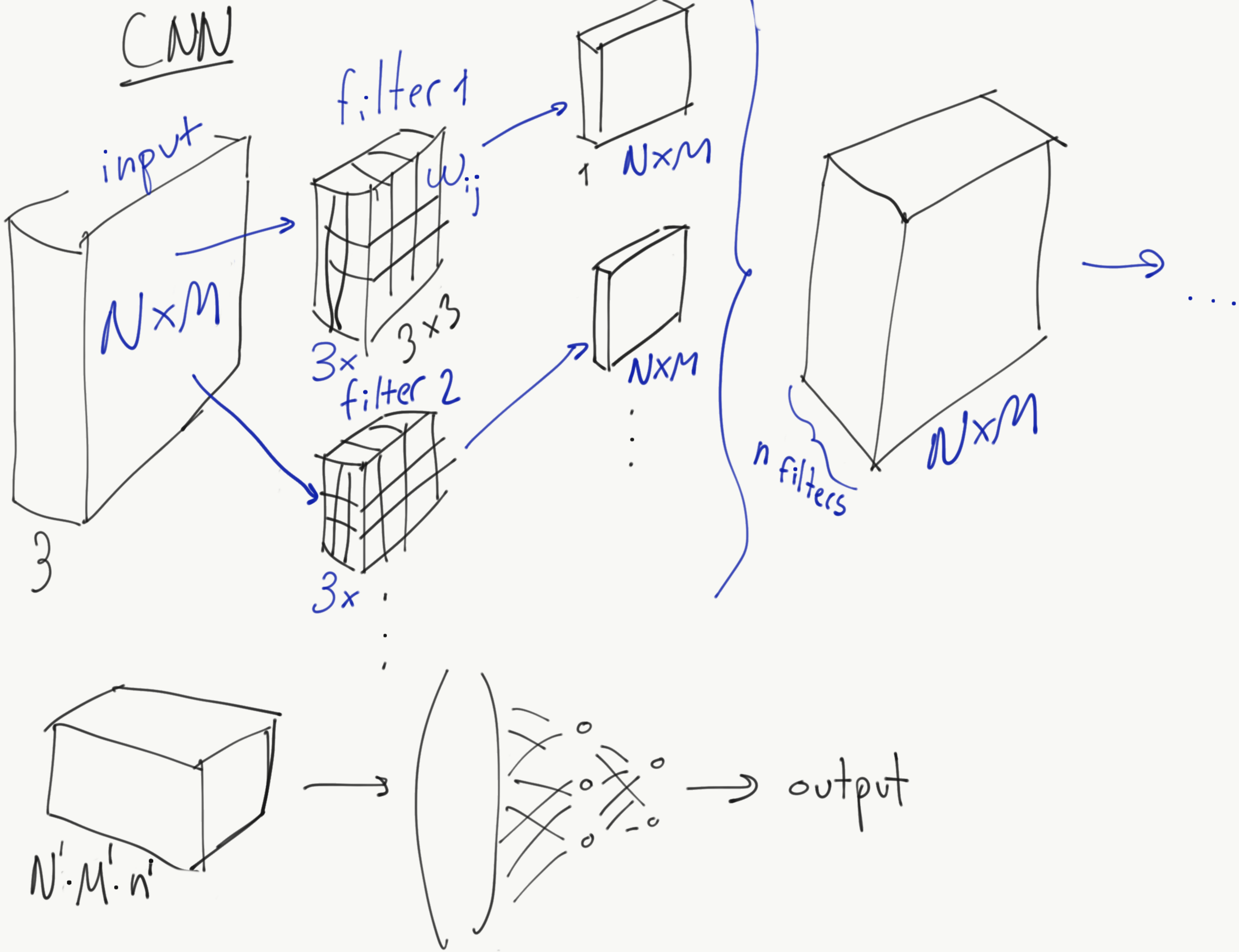
- Fully Connected Networks (FCN)
- Convolutional NN (CNN)
- Generative Models (GM) - training NN to generate new instances
  - AE (Auto Encoders)
  - GAN - Generative Adversarial Networks
- Geometric DL (GDL, e.g. RNN)

Convolutional NN

Around 2010 "fancy" architectures of NN emerged with the sole purpose:
Increase the NN "IQ".
Give the NN the "infra-structure" to build a better language that enable it to better achieve the target.
The first one was the CONVOLUTIONAL NN.



Convolutional NN



What can we (hope to) do with ML?

What can we (hope to) do with ML

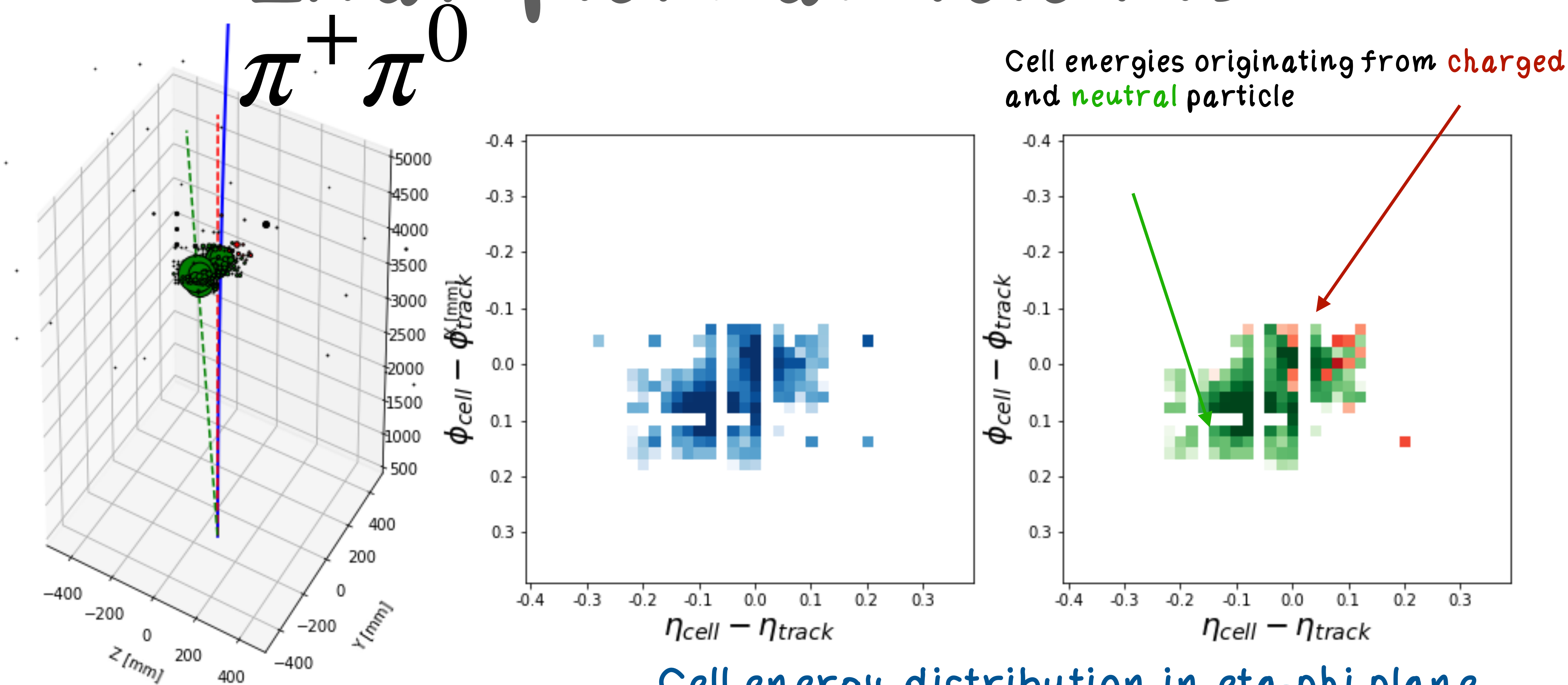
Imaging detection technique for event reconstruction & identification (CNN):

- Detectors become complicated, and might carry information of time, especially in HL environment, A 3D image might become 4D.

- Objects identification is a feature allowing autonomous driving. Its identifying objects in a 2D or 3D pictures that drove the revolution in DNN and influenced tremendously HEP, by giving our dreams food for thought…

- Use CNN on images (e.g. calorimeter response) to identify specific characteristics related to the requested target.
  - ID and measurement of electrons and photons
  - jet properties including b and c tagging, taus and missing energy

Example: Particle Flow



Cell energy distribution in eta-phi plane (centered around track-eta, track-phi)

End to End Deep Learning

- Can we replace the high level features with RAW DATA analysis?
  There is always more information in Low Level DATA.

- However, there is no magic in ML.
  We should design the network in an intelligent way
  (and this still belongs to the human brain…)

What can we (hope to) do with ML

Jet Tagging (CNN, RNN)

- Boosted jets have a substructure which is ideal for image processing with CNN

- A real paradigm change will be when a NN will be applied on RAW data (e.g. tracking, tagging heavy flavours and boosted jets at the trigger level). The ML algorithms do not care about the perigee parameters, they can infer them on passing, from the RAW image.

- HL environment necessities the need for fast and efficient jet taggers. NNs might provide the answer.

Jet Tagging (CNN, RNN)

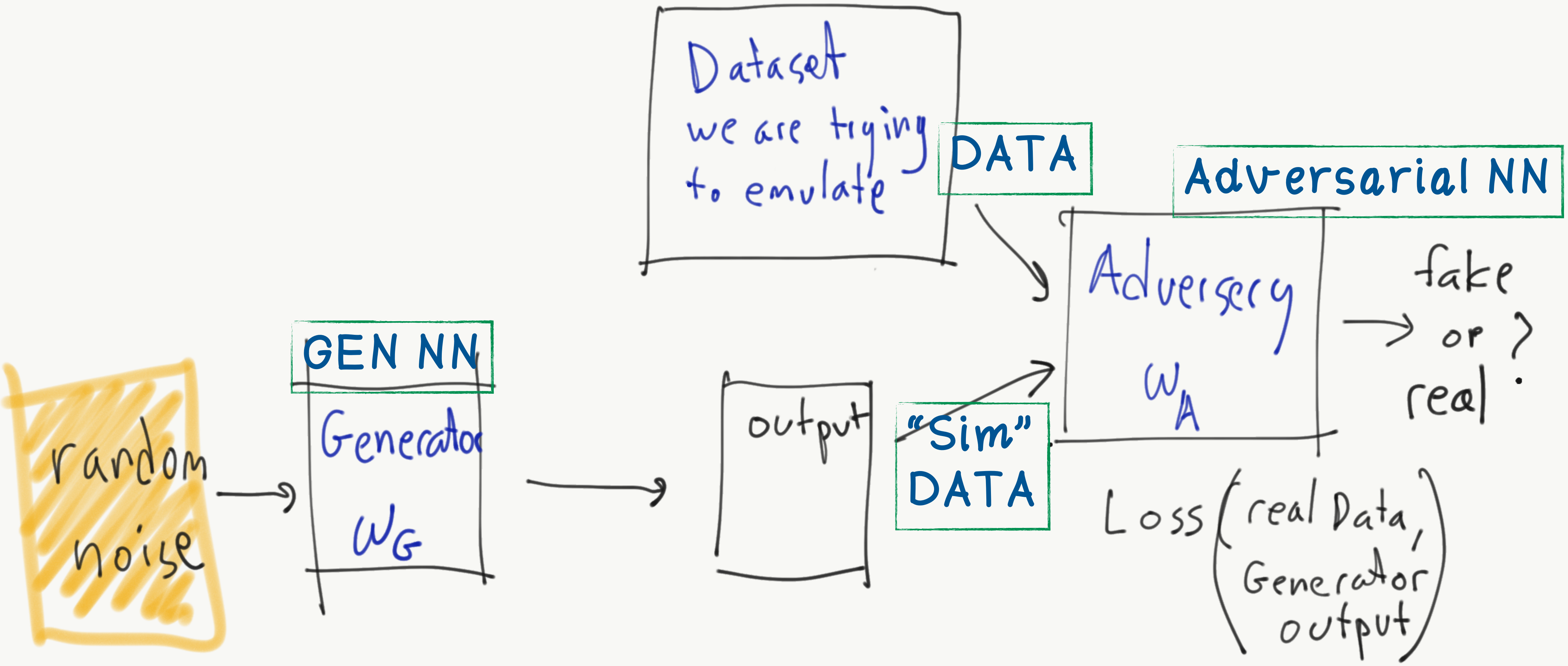
- Imagine a CNN that can do:
  - Multi-class classification
  - Regression for improving energy and mass measurement resolution
  - Event reconstruction with particle flow
  - And all of the above at the detector RAW level

Fast Event Generation (GAN)

- HL-LHC necessitates the need to simulate trillions of events. With O(10 min) to simulate a multi-jets event in GEANT, this becomes a true bottle neck
- Fast simulation exists but at the price of oversimplification
- What we need is to be able to generate the global events in an accurate way without the knowledge of internal details

Fast Event Generation (GAN)

The GEN and the ADVERSARIAL NN are trained against each other, until the classifier cannot distinguish anymore between the DATA events and those provided by the generator (SIM).



train in cycle:

1. modify w_A to decrease loss
2. modify w_G to increase loss

What can we (hope to) do with ML

Fast Event Generation (GAN)

- Use neural networks to generate events similar to those used to train a network .
- This will allow faster detector simulations while preserving the accuracy of a full simulation.
- GANs offer an alternative to simulation, with an order of magnitude improvement in the simulation speed, but yet not accurate enough. There is still a long way to go till GANs are effectively used and are able to reproduce the interactions of particles with material and "simulate" a real detector.
- GEANT is so slow that its hard to imagine one can keep on using it for HL.
- R&D is required to improve the GANs, but its a highly promising avenue to pursue.

Reducing systematics (GAN)

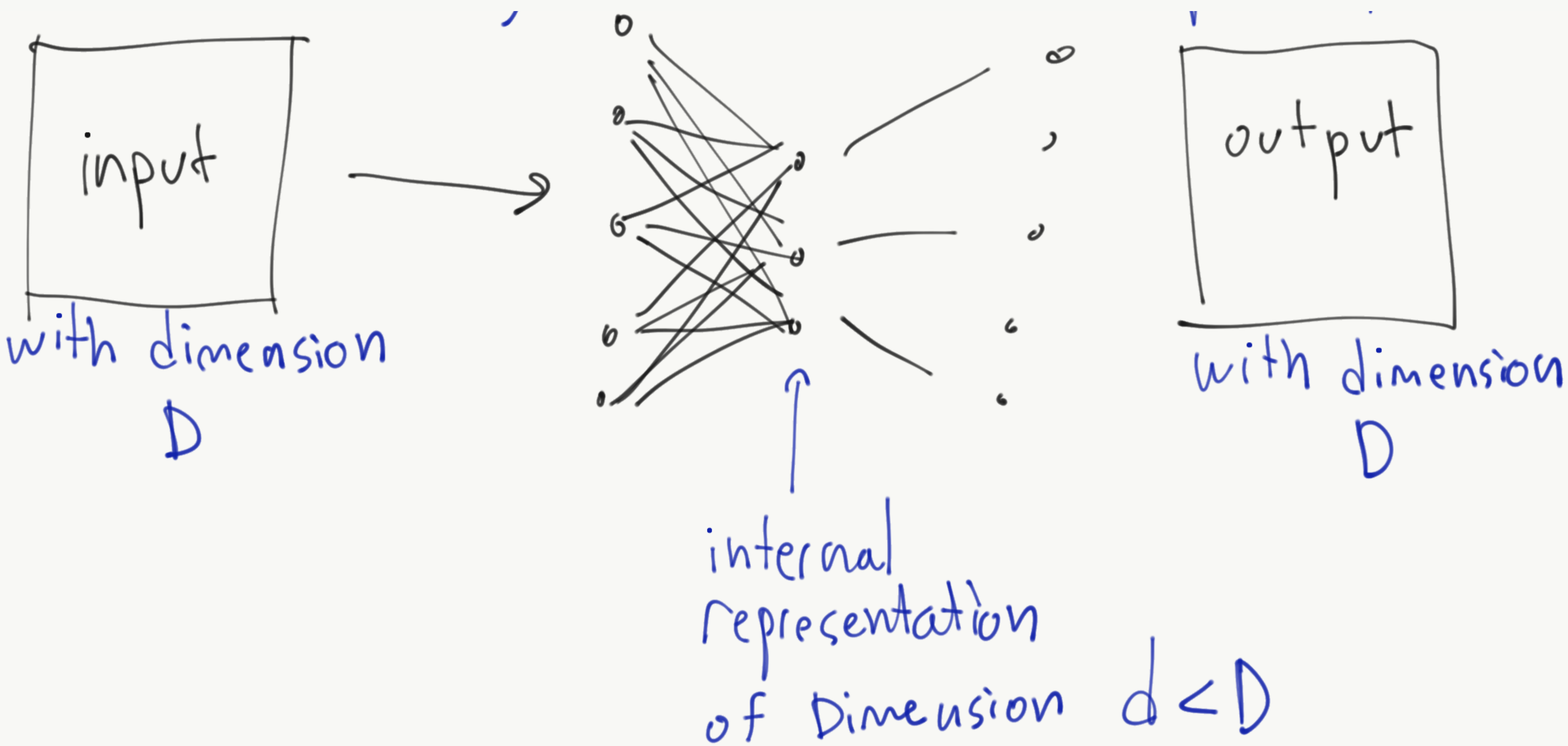
- One can use the adversarial component of the GAN to train the network to be insensitive to a specific systematics at the price of reduced (classifier) performance

Search for New Physics (Anomalies with Auto Encoder)

Train a simple neural network to "encode" the data in a representation with lower dimension, and decode from it an image that is supposed to match the input

Train the network to approximately
reproduce the input

LOSS is the "distance" between the decoded output and the input



Search for New Physics (Anomalies with Auto Encoder)

- Traditional searches are tuned to find evidence for or against a given NP model.
- Use Auto Encoders to teach the network how the SM "looks like" and then identify anomalies in the REAL DATA with respect to the SM.
- Auto encoders work!
- BUT, the problem is at the trigger level. We might throw away NP events without even knowing.
  Can AEs be used to prevent this?

NP mining at the trigger level

- There are whole classes of events, for example beauty and charm hadrons or low-mass dark matter signatures, which are so abundant that it is not affordable to store all of the events for later analysis.
- In order to fully exploit the physics reach of the LHC, it will increasingly be necessary to perform more of the data analysis in real-time.
- ML may be the only hope of performing real-time reconstruction that enables real-time analysis.
- The hope is that ML will be able to fully do tracking, vertexing. and jets analysis at early stages allowing to lower for example, jet trigger thresholds.

NP mining at the trigger level

- Replace the scouting and/or hotline streams or TLA at the trigger level by an algorithm that will not loose "unknown" outlier events but save them for later inspection

  - This might be done by unsupervised and semi-supervised encoders trained to associate an event topology to a known physics process.
    Unspecified anomalous events, as those coming from new physics, would be saved as
    "unlabelled events" (rather than being rejected by the trigger) for offline analysis.
    Different algorithms of this kind could be developed for different topologies

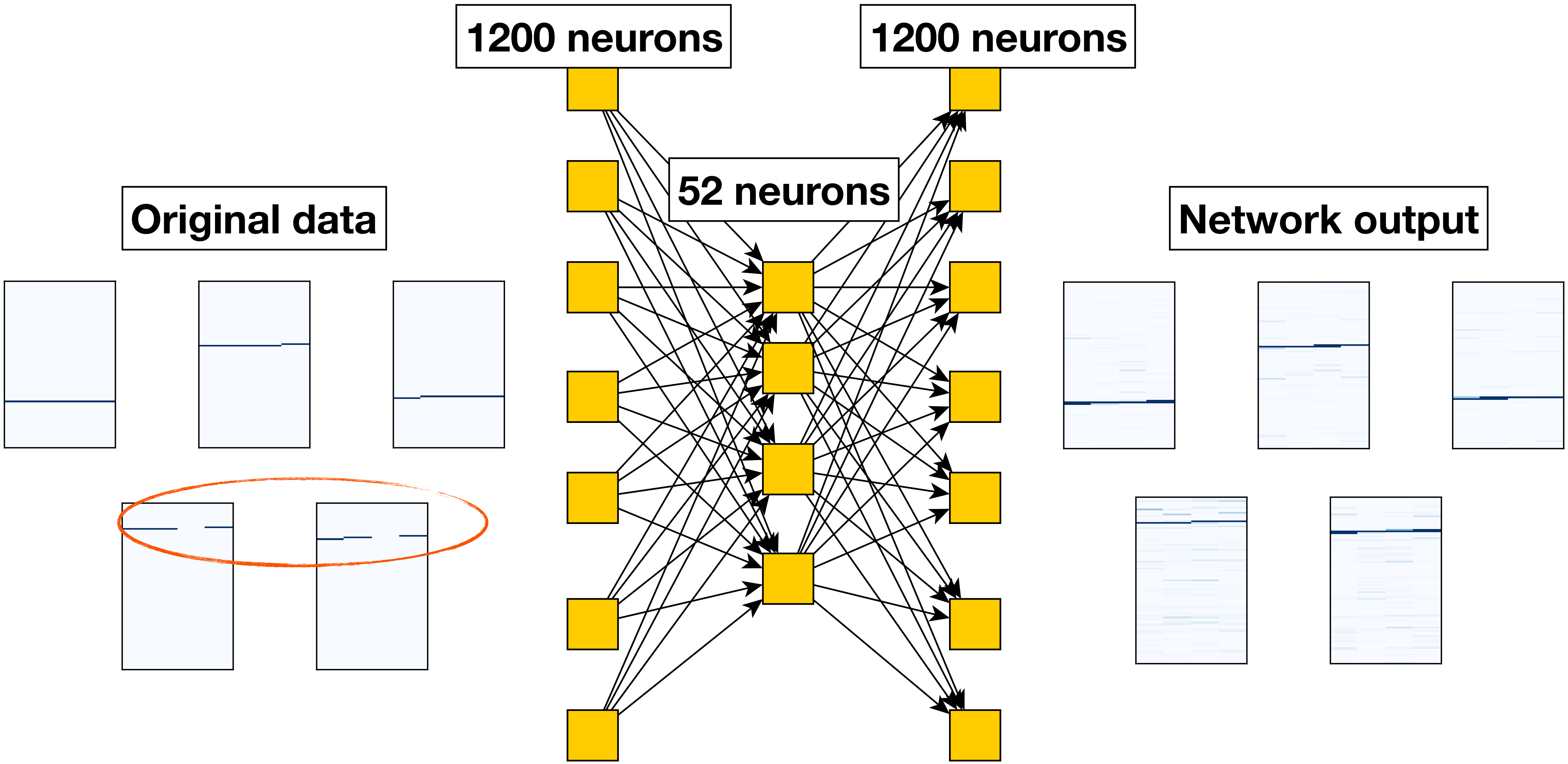
  - Alternatively, auto-encoders can do a more efficient job of labelling Physics events allowing the inclusion of more classes in the trigger streams

DAQ Monitoring

- Develop and deploy an anomaly detection application that identifies when the collected data deviate from a specified normal behavior. This application will assist the DQM shifter and could eventually replace him.

- This can be done for specific Detectors or for the DATA stream itself, certifying DATA for Physics Analysis (GOOD RUNS)

Autoencoder



Other Applications of ML

- Anomaly detection is not only about a 750 GeV scalar that nobody has invited… Its also about a detector that does not exhibit a well expected behaviour due to some Hardware or maintenance failure. Here ML techniques can be very useful.

- ML can be used to predict the popularity of a dataset from dataset usage and reduce disk resource utilizations, improving analysis time turnover…

- ML can detect anomalies in network traffic, predict network congestion, detect bugs etc…..

- Theorists Paradise:
  CAN PDF with its uncertainties be extracted using NNs?
  The NNPDF collaboration determines the structure of the proton using contemporary methods of artificial intelligence.

Assigning Uncertainty to ML

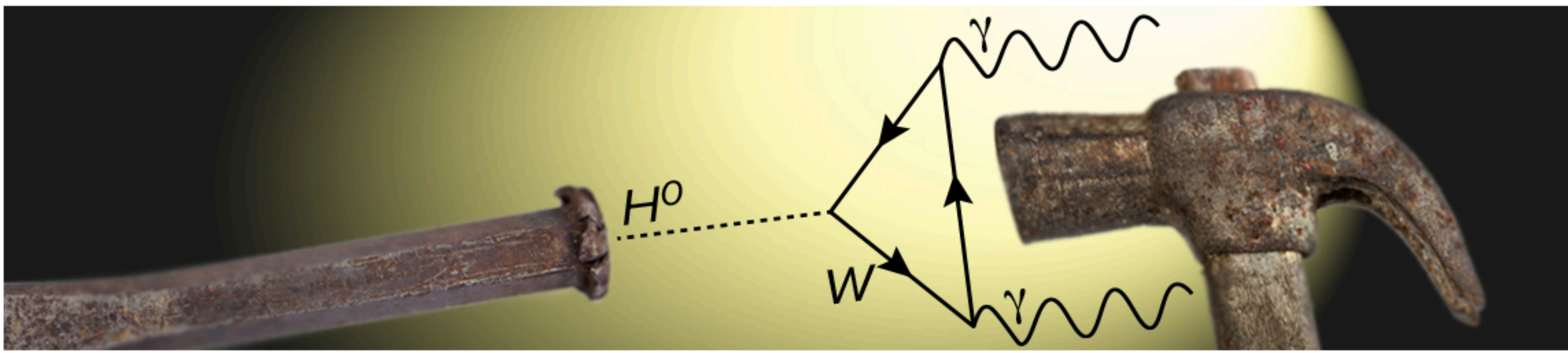
- No real advance in ML techniques for Data analysis can be made without learning how to assign uncertainty to the algorithms. This is still an ongoing subject without a well defined answer.

ML Training

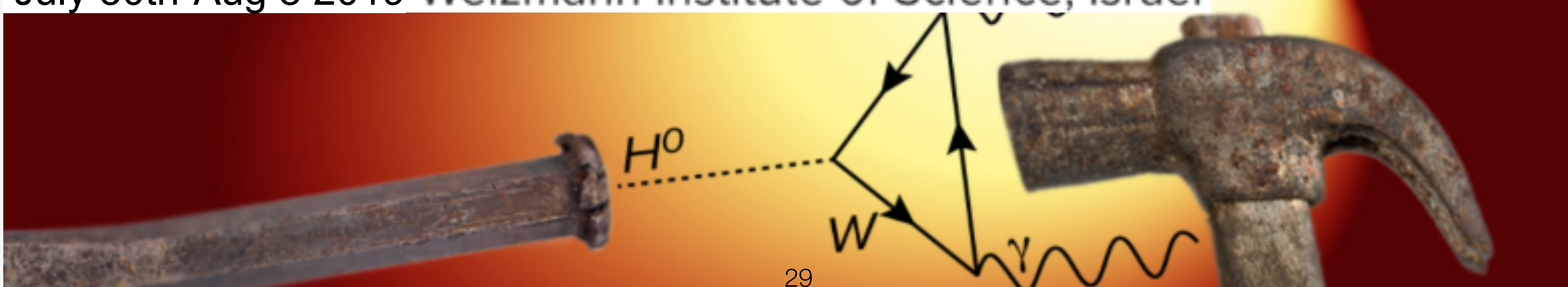
- Physicists need to learn how to pause the right input and target and consult the computer DATA scientists about the best architectures to solve their problem

Hammers & Nails - Machine Learning & HEP

July 19-28, 2017 | Weizmann Institute of Science, Israel



July 30th-Aug 8 2019 Weizmann Institute of Science, Israel



What can we (hope to) do with ML
The MECCA of ML in HEP

- can ML eventually run the experiment using Reinforced Learning
- Reinforced algorithms based on Deep Learning DAQ Monitoring, Hardware Failure detection auto encoders can make recommendation for best operation of the detector (or even take the decisions….).
- In short WE CAN GO HOME

THANK YOU FOR YOUR PATIENCE