

Machine Learning Course 4

Advanced ML for HEP

(more on unsupervised learning, anomaly detection, clustering
specific HEP challenges in Ana Peixoto's courses next week)



David Rousseau
IJCLab-Orsay

david.rousseau@in2p3.fr

@dhpmrou

CHACAL, Johannesburg, Jan 2024



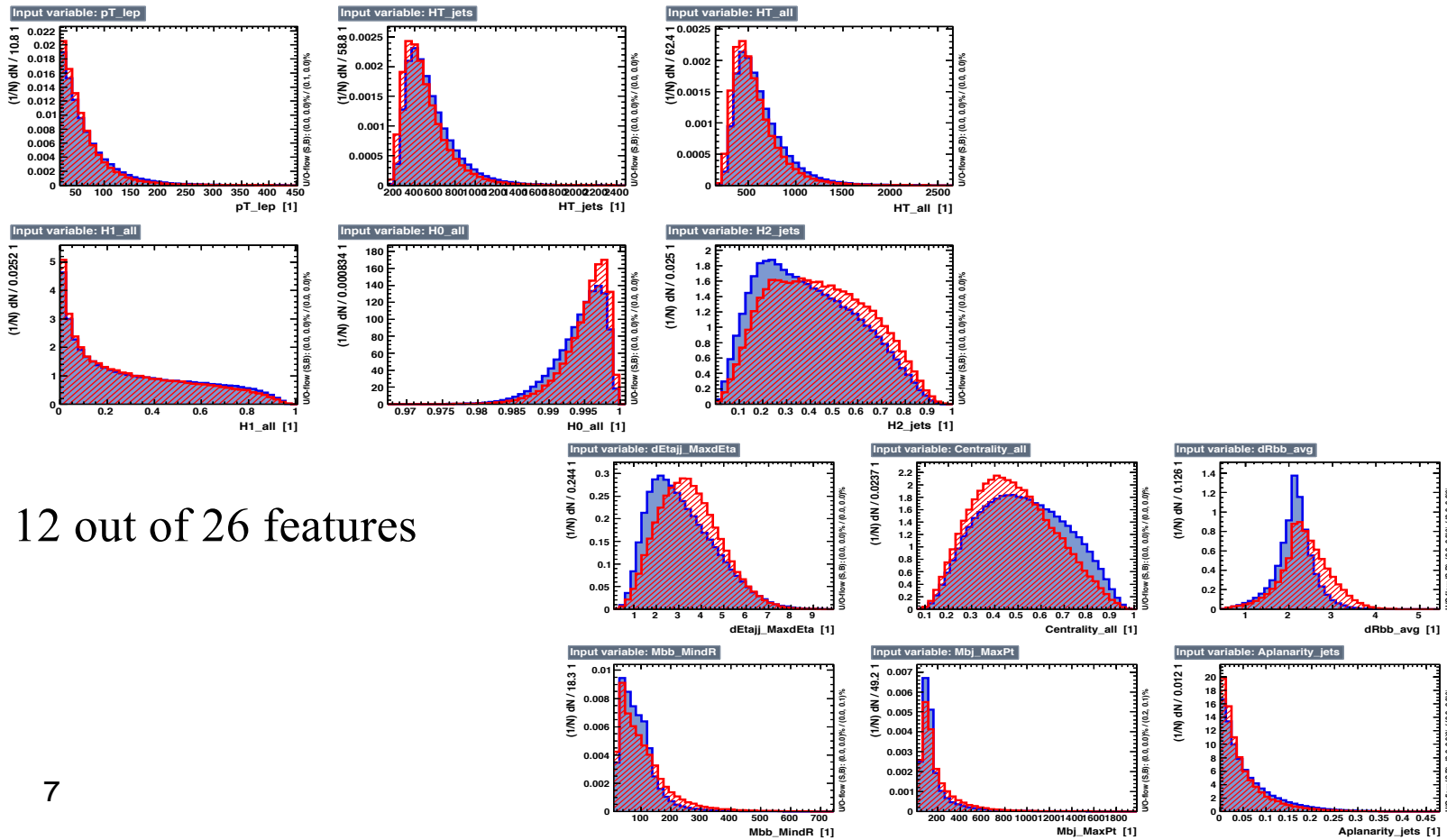
Feature Selection



BB.

Feature Selection

Example of an ATLAS experiment analysis, [see transp](#)



Feature selection techniques

- > Iterative addition: start with $n=1$, take best training of all n options. Then take best option of adding one more from remaining $n-1$ variables, etc. Ranking complexity scales as $O(n^3)$.
- > Iterative removal: start with training on all variables and remove iteratively remove the one that degrades the performance the least, scales as $O(n^3)$.
 - > Hypothesis: better consideration of variables that only add to performance in combination with others.
- > Correlation based: rank the variables based on their correlation to the BDT score computed with all variables. Computationally cheap, scales as $O(n)$.
- > BDT selection frequency 'TMVA ranking': train once on all variables, rank by how often a variables provided the optimal decision in the BDT, scales as $O(n)$.
- > Separation based: rank by overlap of signal vs background shapes. Only method that establishes ranking without performing any training.
- > Random choice: serves as reference, use a random subset of the variables. Repeat and average over 1000 trials.

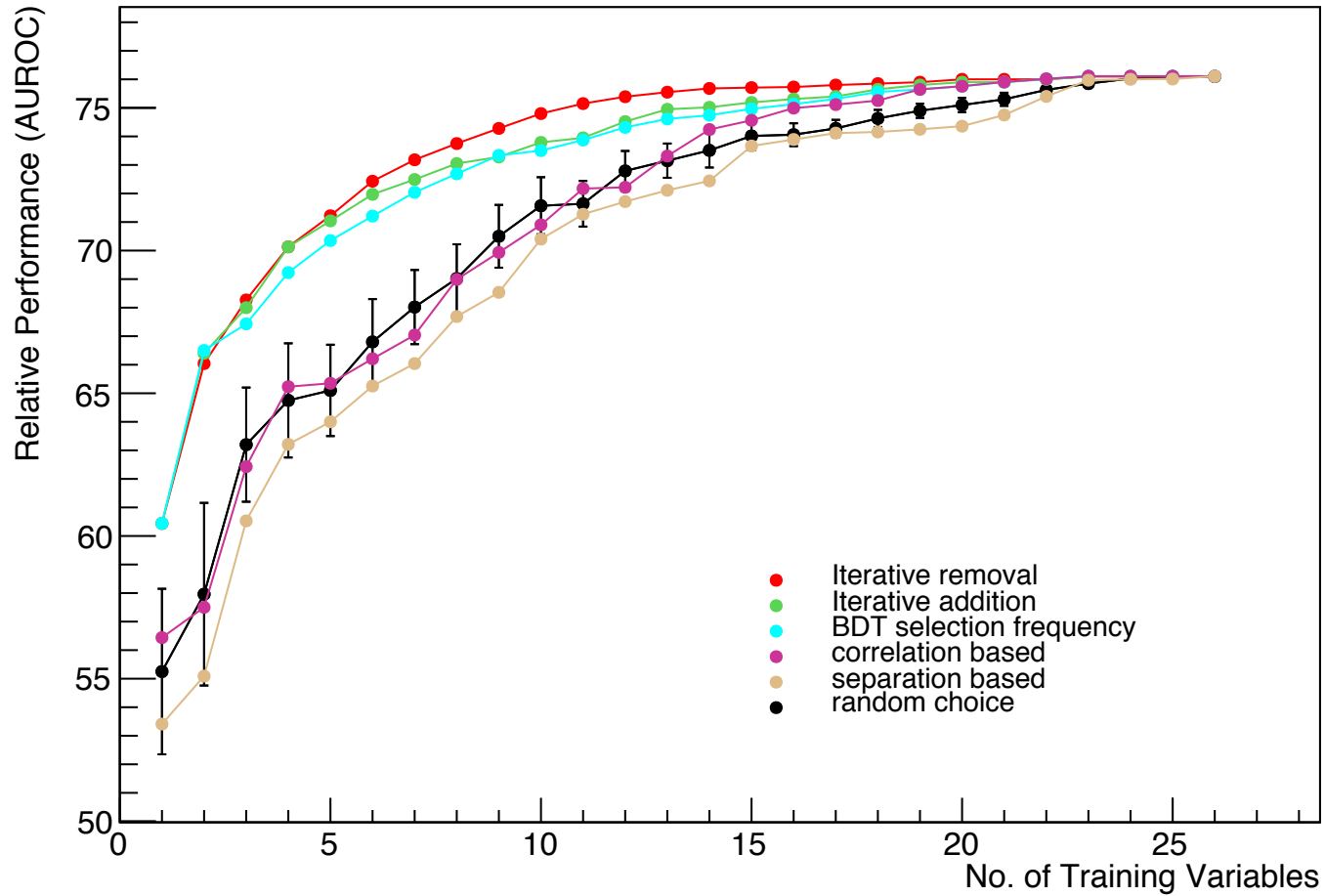
Key role permutation importance

Here performance is measured as integral of the ROC curve.



Results

AUC vs No. of Variables



Conclusion : iterative
removal based on
Permutation
Importance

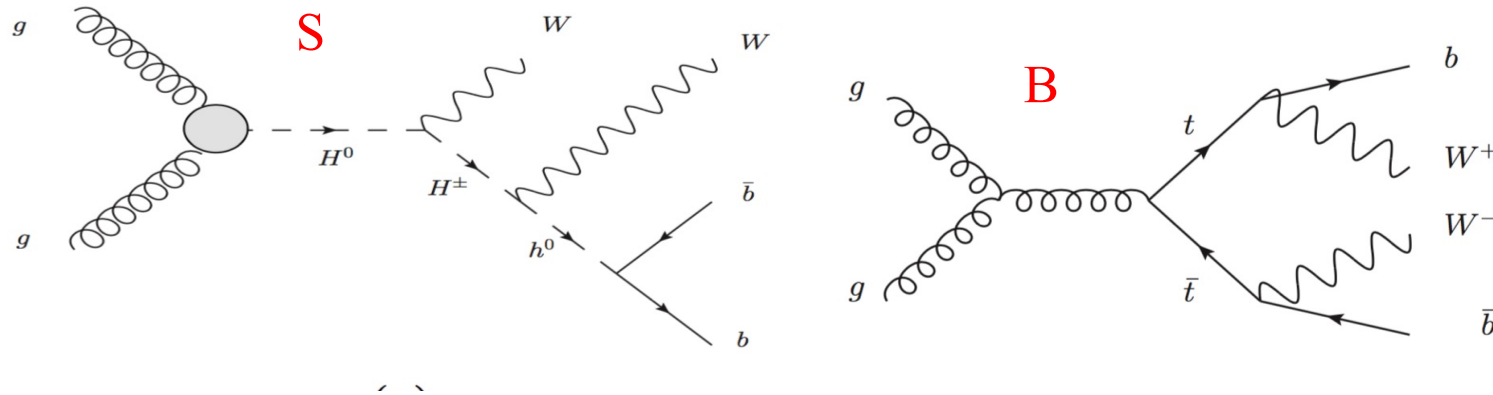


Deep Learning for Analysis



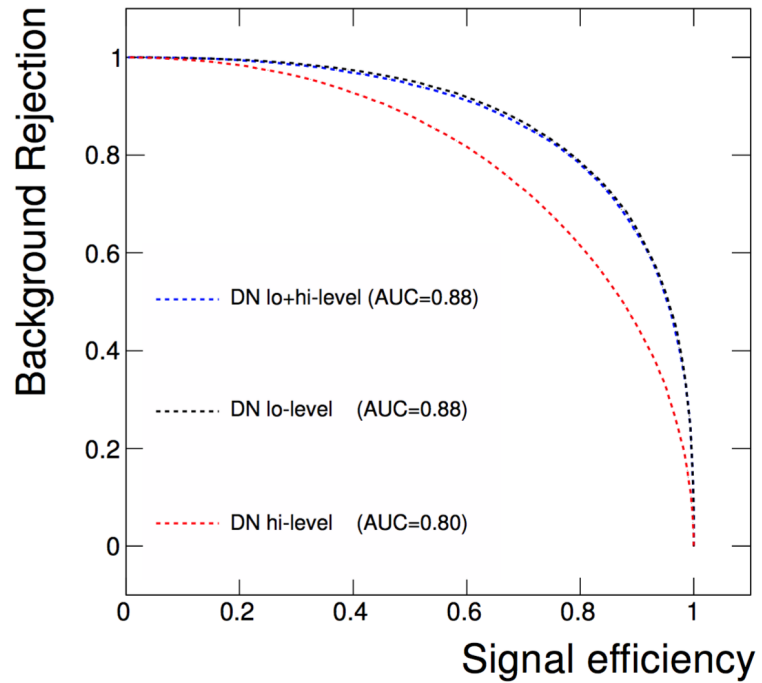
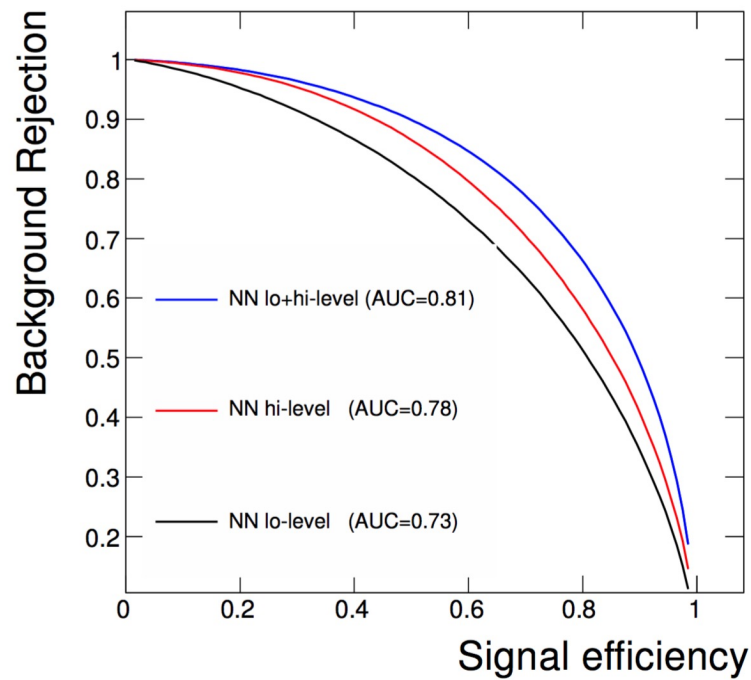
Deep Learning for Analysis

1402.4735 Baldi, Sadowski, Whiteson



- ❑ MSSM at LHC : $H^0 \rightarrow WWbb$ vs $t\bar{t} \rightarrow WWbb$
- ❑ Low level variables:
 - 3-momentum vectors
- ❑ High level variables:
 - Pair-wise invariant masses
- ❑ Compare shallow NN (1 hidden layer) with (not so) deep NN (3 hidden layers)

Results

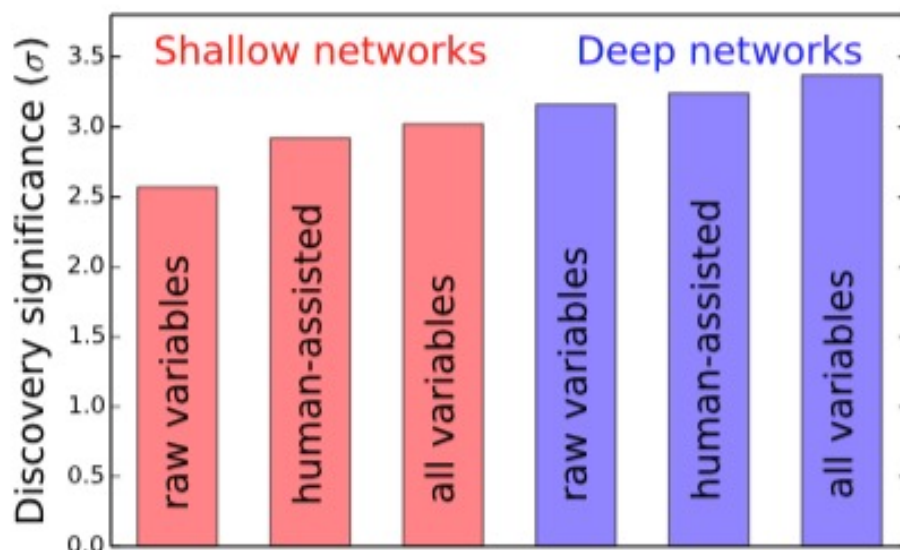


- ❑ Deep NN outperforms NN, and does not need high level variables
- ❑ DNN learns the physics ???

Deep learning for analysis (2)

1410.3469 Baldi Sadowski Whiteson

- H tautau analysis at LHC: $H \rightarrow \text{tautau}$ vs $Z \rightarrow \text{tautau}$
 - Low level variables (4-momenta)
 - High level variables (transverse mass, delta R, centrality, jet variables, etc...)

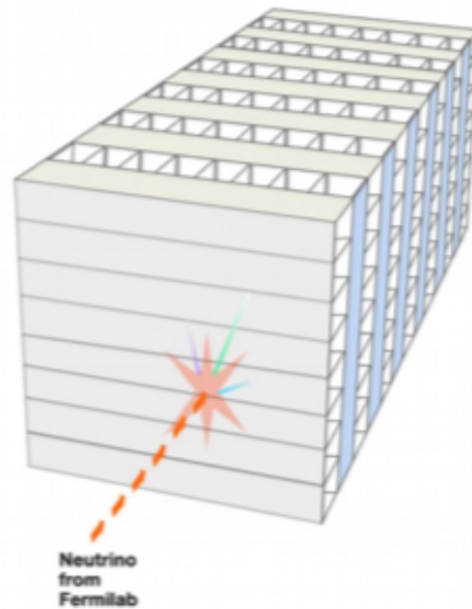


- Here, the DNN improves on NN but ...
- ...DNN still needs high level features
- Both analyses with Delphes fast simulation $\sim 100\text{M}$ events used for training (very unusual $\gg 100^*$ full G4 simulation in ATLAS)
- No convincing proof so far deep dense NN better than shallow NN or BDT, in realistic conditions

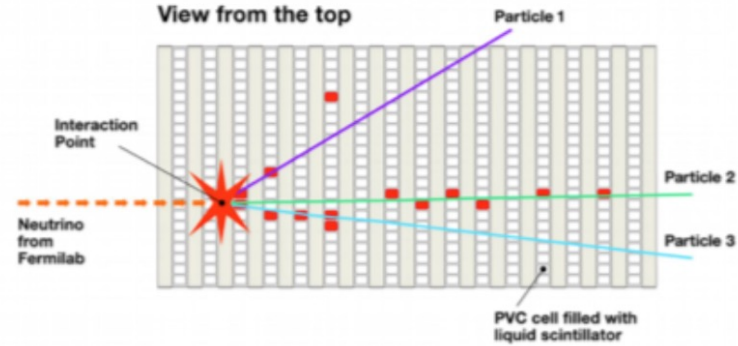
Deep Learning success : NOVA



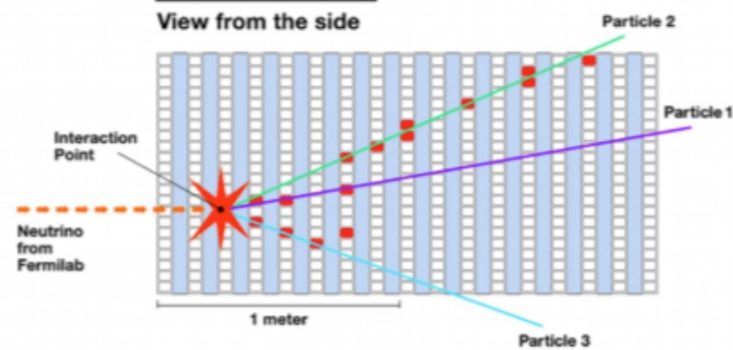
3D schematic of NOvA particle detector



View from the top



View from the side



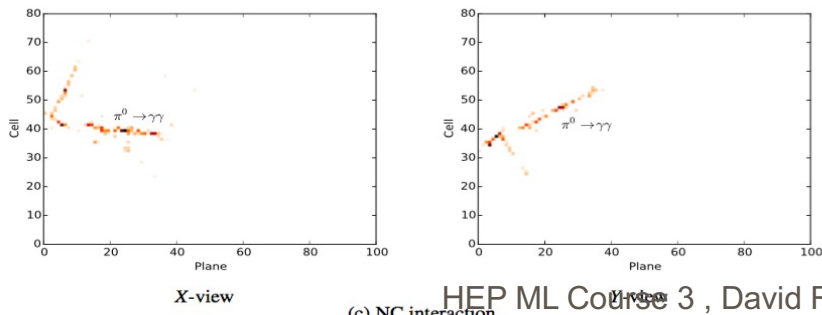
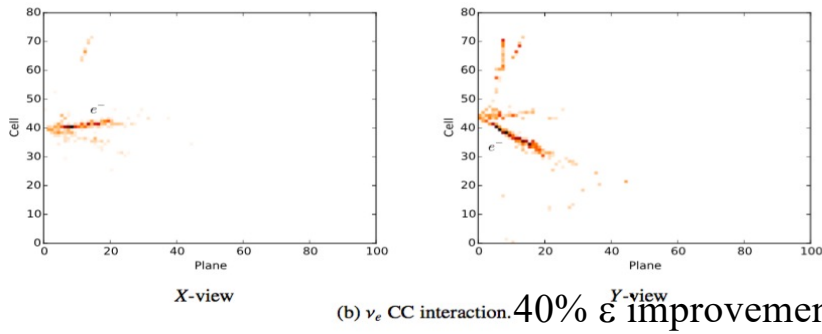
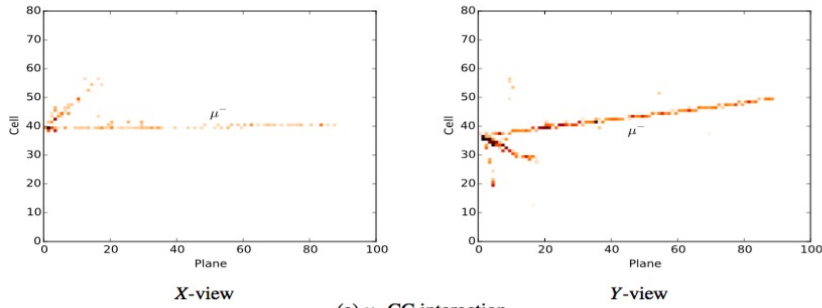
e
b
ing



HEP ML Course 3 , David Rousseau, Jan 2024, CHACAL

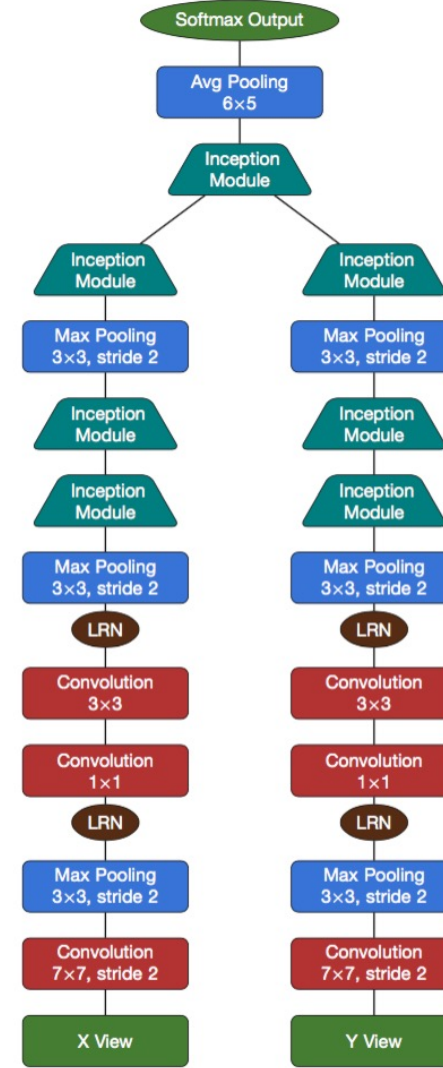
Nova (2)

arXiv 1604.01444 Aurisano et al



Neutrino interaction classification
Using Convolutional Neural Network (GoogLeNet)

Actually used in physics results 1703.03328 and 1706.04592



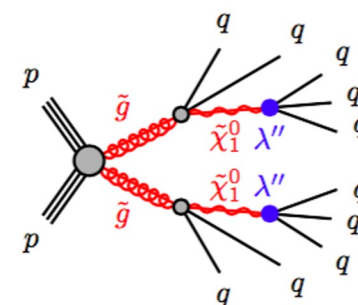
End to end Learning



End to end learning

Bhimji et al, 1711.03573

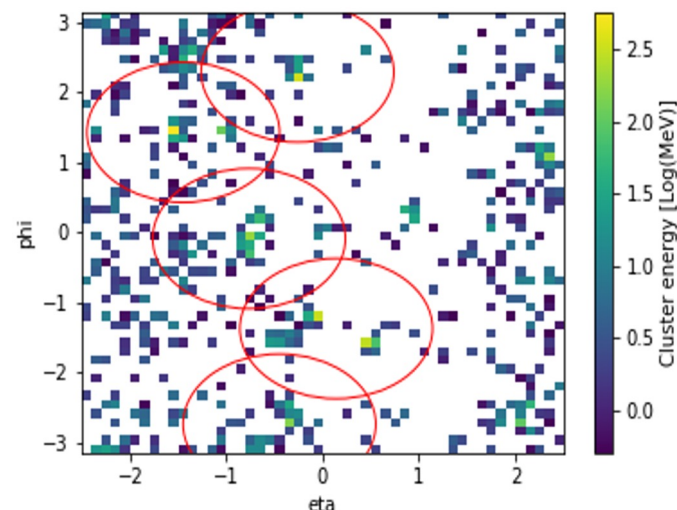
- ❑ Train directly for signal on « raw » event ?
- ❑ Start from RPV Susy search
ATLAS-CONF-2016-057
- ❑ Fast Simulated events with Delphes



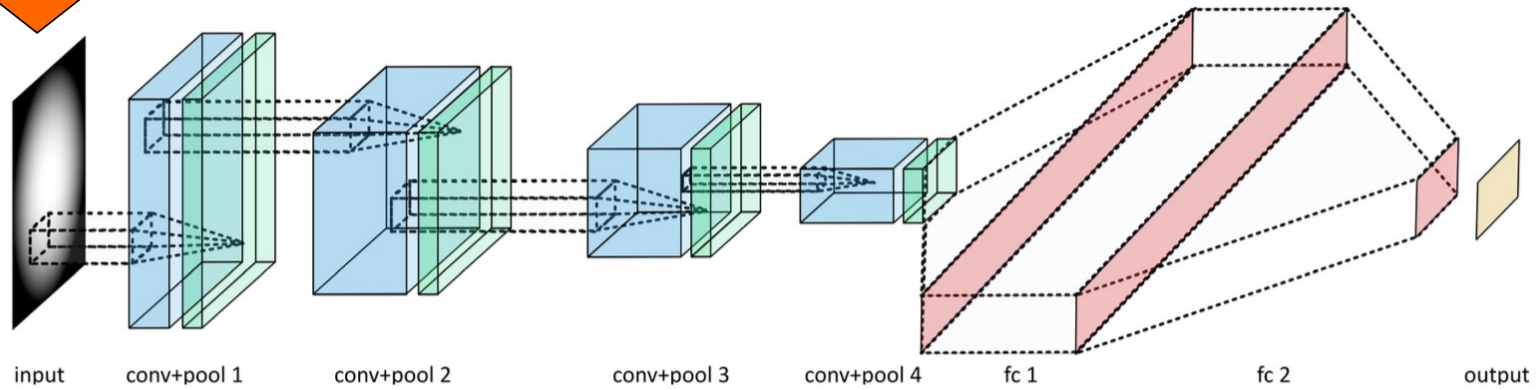
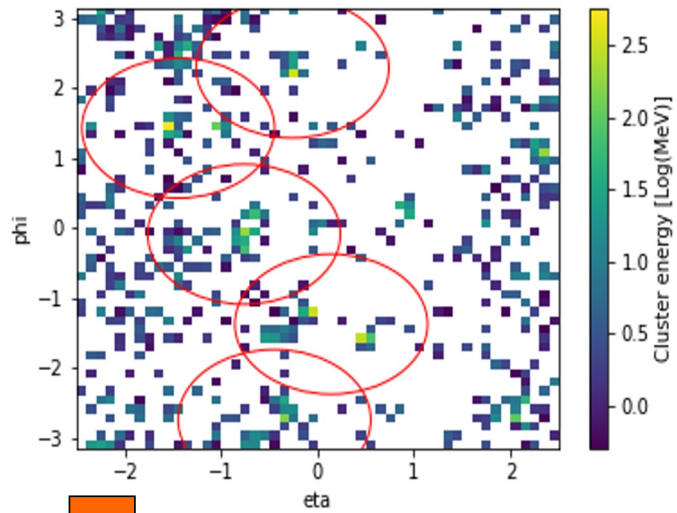
(b) gluino cascade decay

- ❑ Project energies on 64x64 $\eta \times \phi$ grid
- ❑ Compare with usual jet Reconstruction and physics Analysis variables such as:

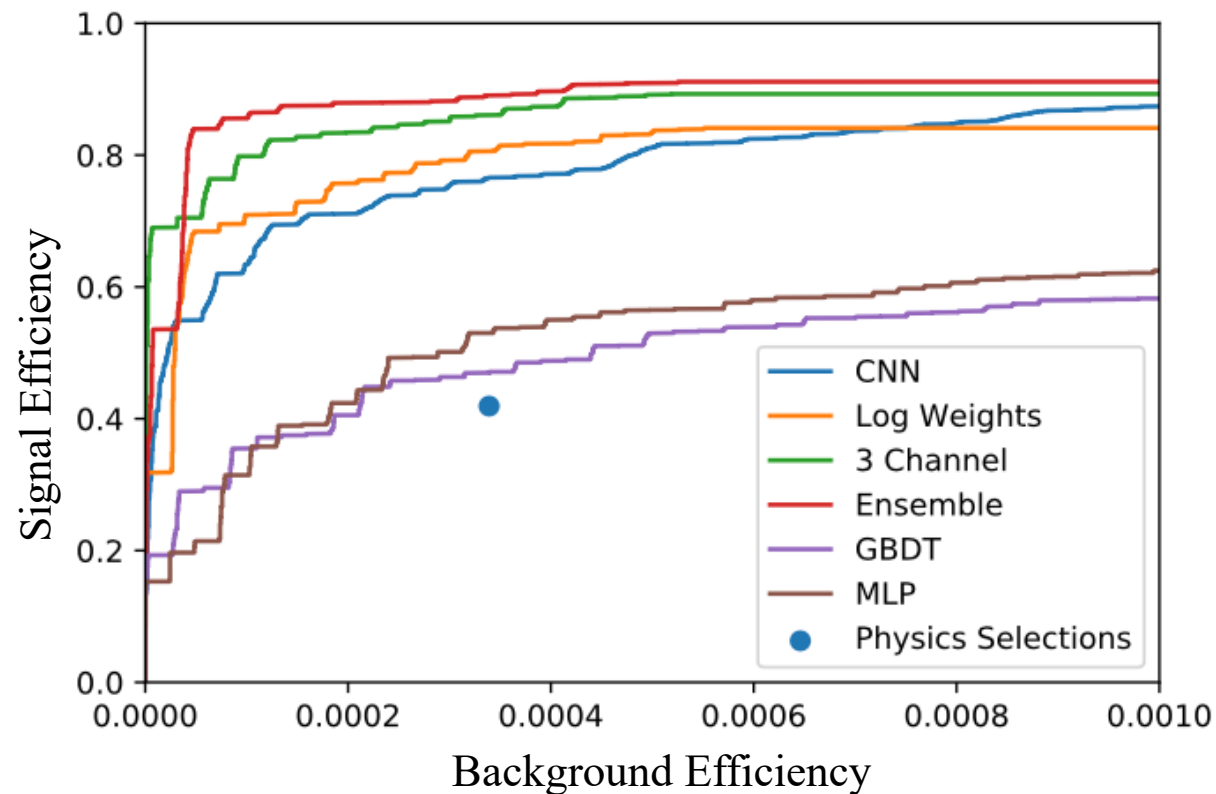
$$M_J^\Sigma = \sum_{\substack{p_T > 200 \text{ GeV} \\ |\eta| \leq 2.0}}^4 m^{\text{jet}}$$



End to end learning (2)



End to end learning (3)



- ❑ >x2 gain over BDT/shallow network using physics variable and 5 leading jet 4-momenta
- ❑ →CNN extract information from energy grid which is lost in the jets ?
- ❑ Not sure, they should compare to applying DL on the jets

Systematics Aware Training

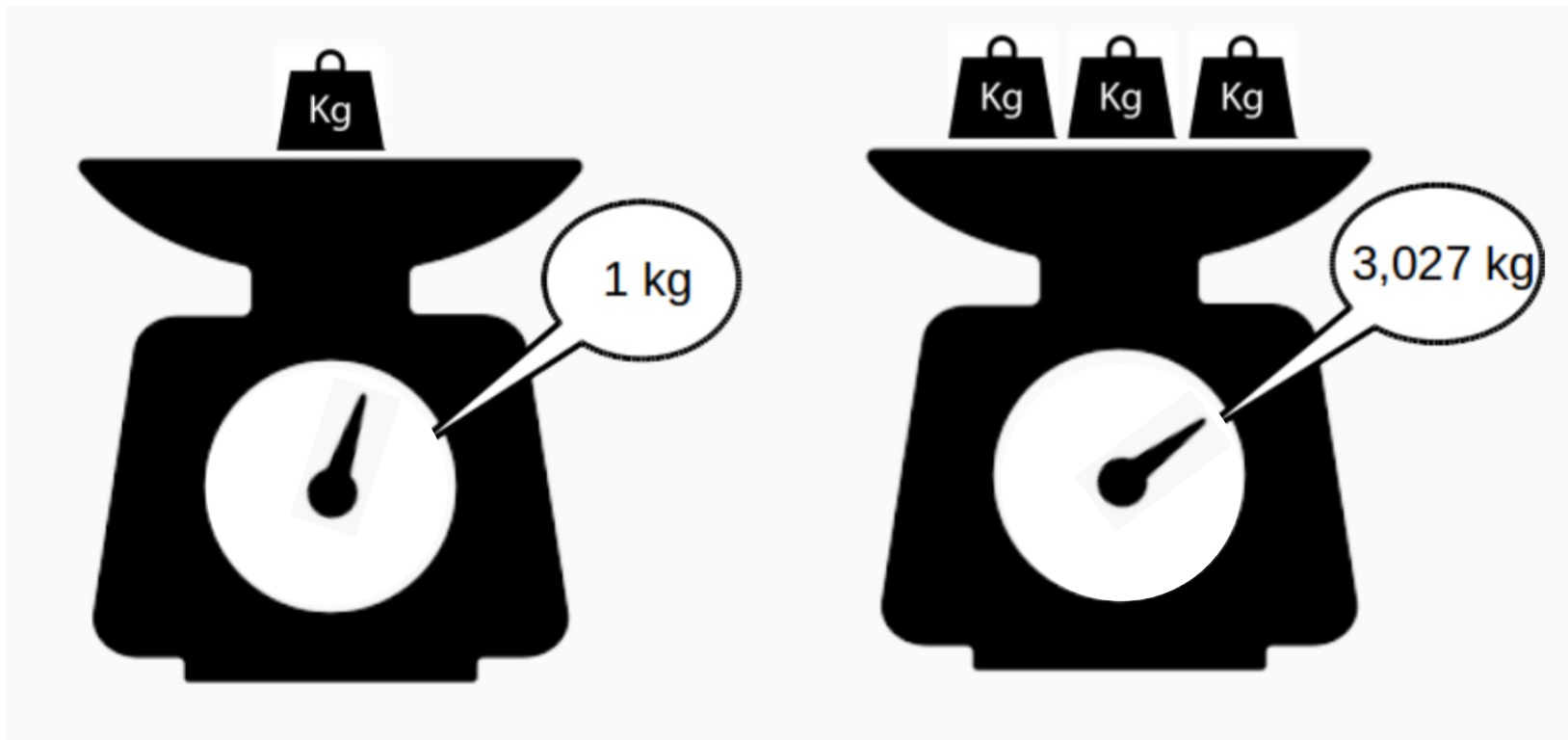


Systematics-aware training

See Victor Estrade CHEP 2018

- ❑ Our experimental measurement papers typically ends with
 - measurement = $m \pm \sigma(\text{stat}) \pm \sigma(\text{syst})$
 - $\sigma(\text{syst})$ systematic uncertainty : known unknowns, unknown unknowns...
- ❑ Name of the game is to minimize quadratic sum of :
$$\sigma(\text{stat}) \pm \sigma(\text{syst})$$
- ❑ ML techniques used so far to minimise $\sigma(\text{stat})$
- ❑ Impact of ML on $\sigma(\text{syst})$ or even better global optimisation of $\sigma(\text{stat}) \pm \sigma(\text{syst})$ is an open problem
- ❑ Worrying about $\sigma(\text{syst})$ untypical of ML in industry (... until recently *fake news*)
- ❑ However, a hot topic in ML in industry: *transfer learning*
- ❑ E.g. : train image labelling on a image dataset, apply on new images (different luminosity, focus, angle etc...)
- ❑ For HEP : we train with Signal and Background which are not the real one (MC, control regions, etc...) → source of systematics

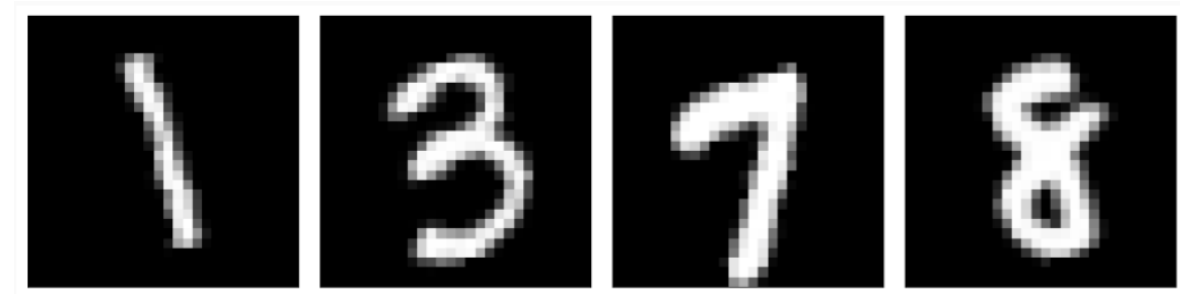
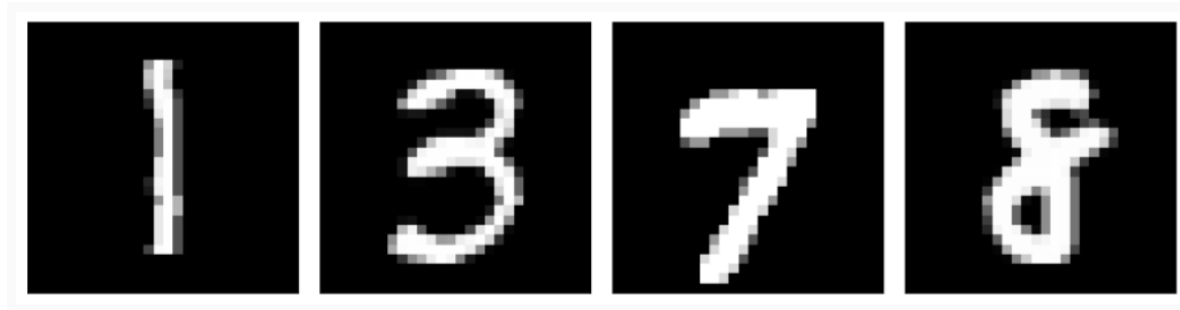
Experimental bias



Systematical effect



Example of impact of the angle on handwritten digits



Example mHiggs measurement

About 500 terms for the systematical uncertainty

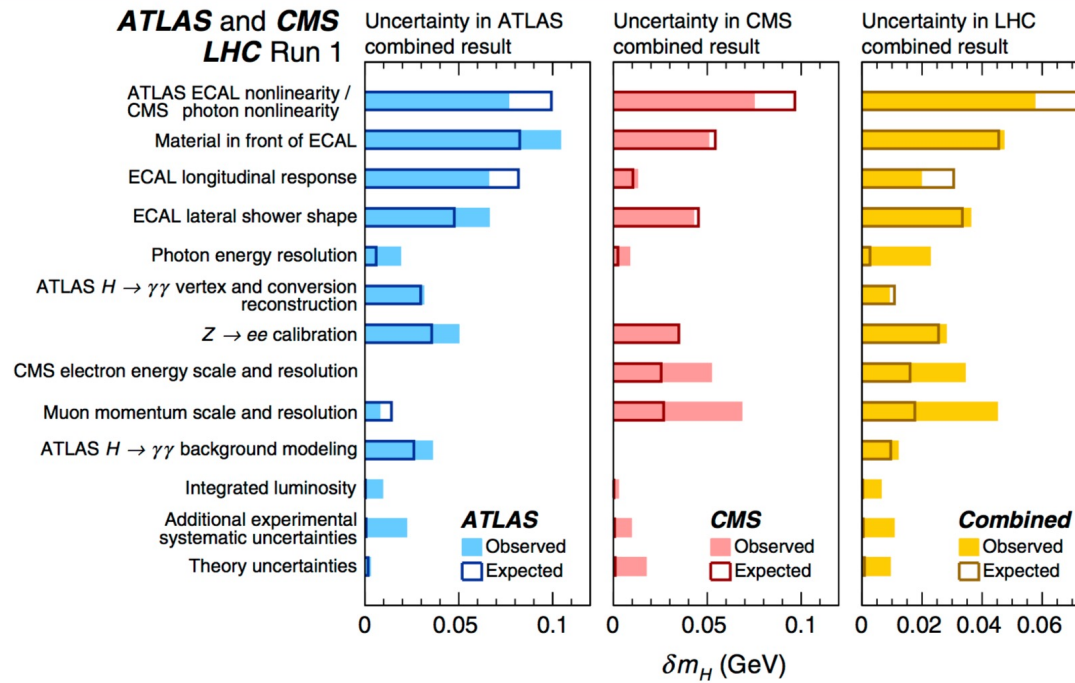
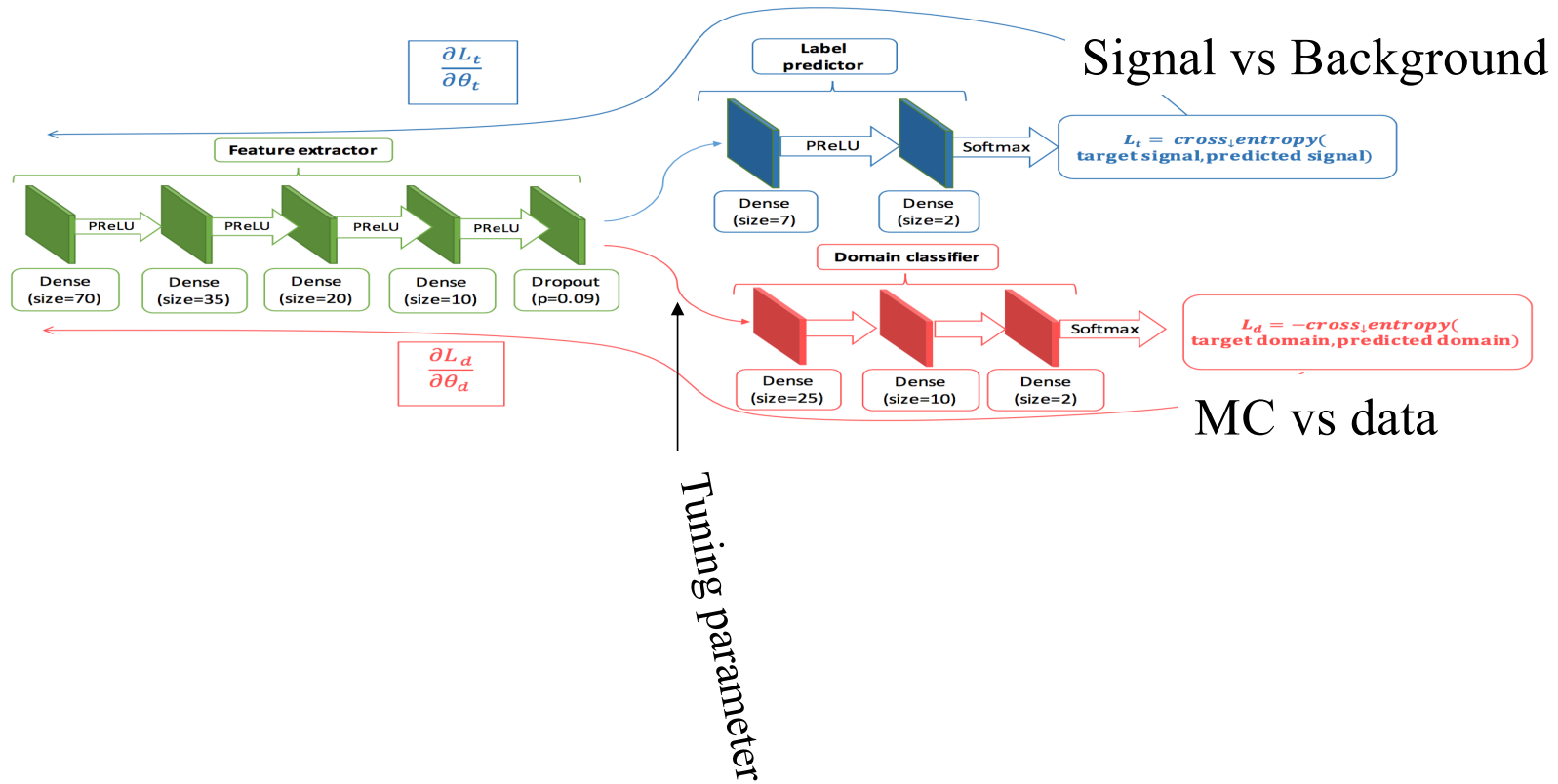


FIG. 3 (color online). The impacts δm_H (see text) of the nuisance parameter groups in Table I on the ATLAS (left), CMS (center), and combined (right) mass measurement uncertainty. The observed (expected) results are shown by the solid (empty) bars.

Syst Aware Training: adversarial

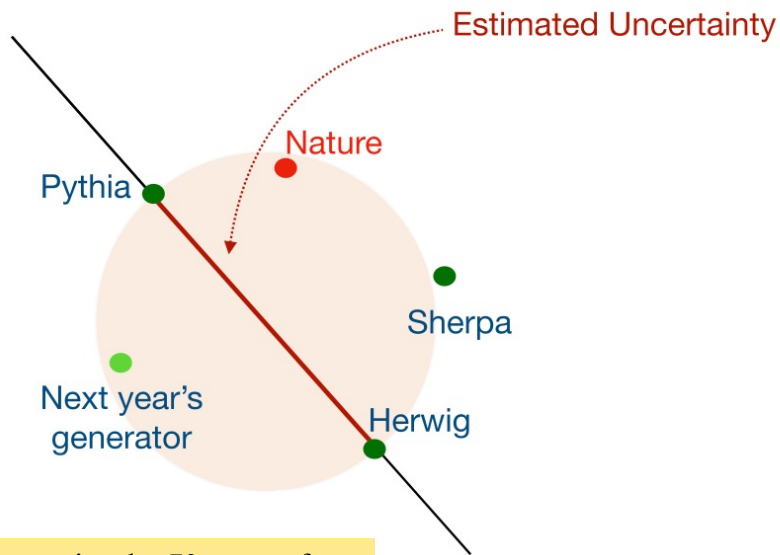
Inspired from 1505.07818 Ganin et al :



Not conclusive yet

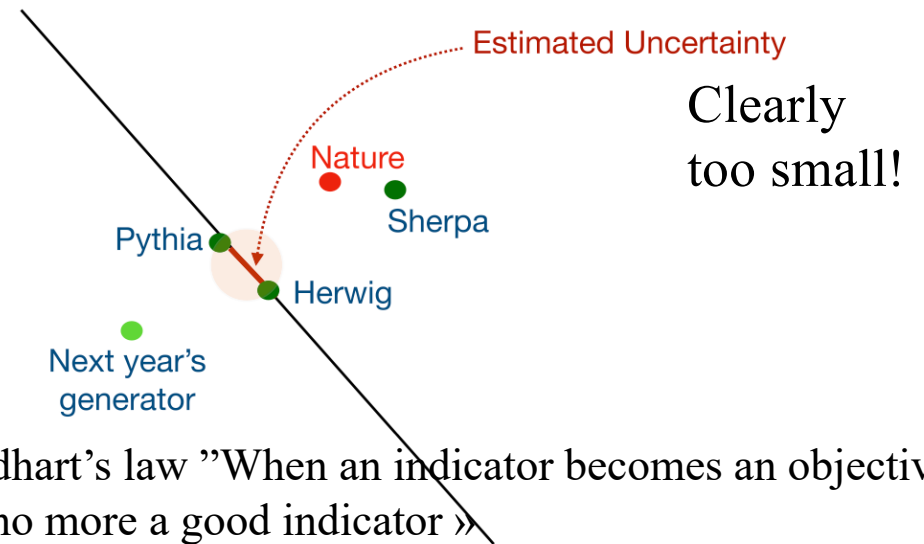
Ghosh & Nachman EPJC 82 46 (2022)

Without Decorrelation



Constraints by 70 years of Particle Physics measurement

With Decorrelation



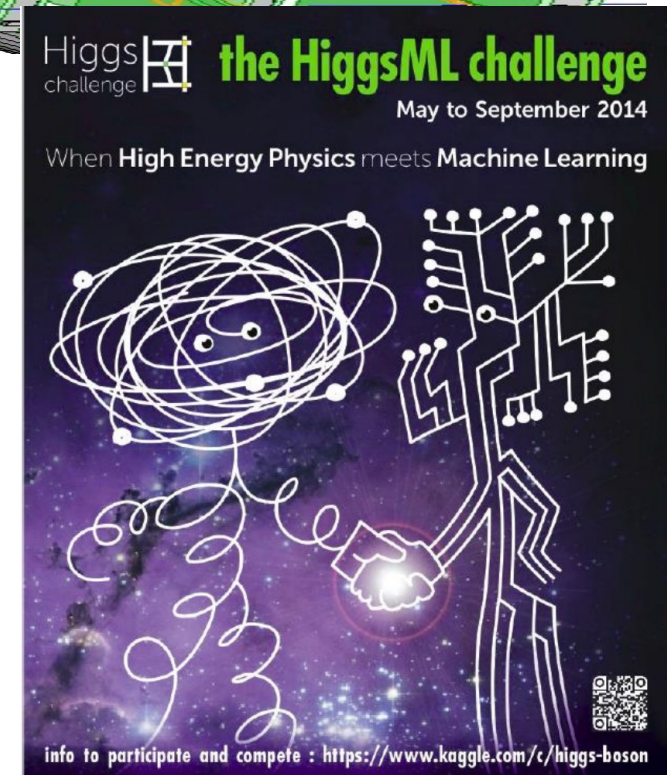
Clearly too small!

Goodhart's law "When an indicator becomes an objective, it is no more a good indicator »"

Hot topic!

Fair Universe: HiggsML Uncertainty Challenge

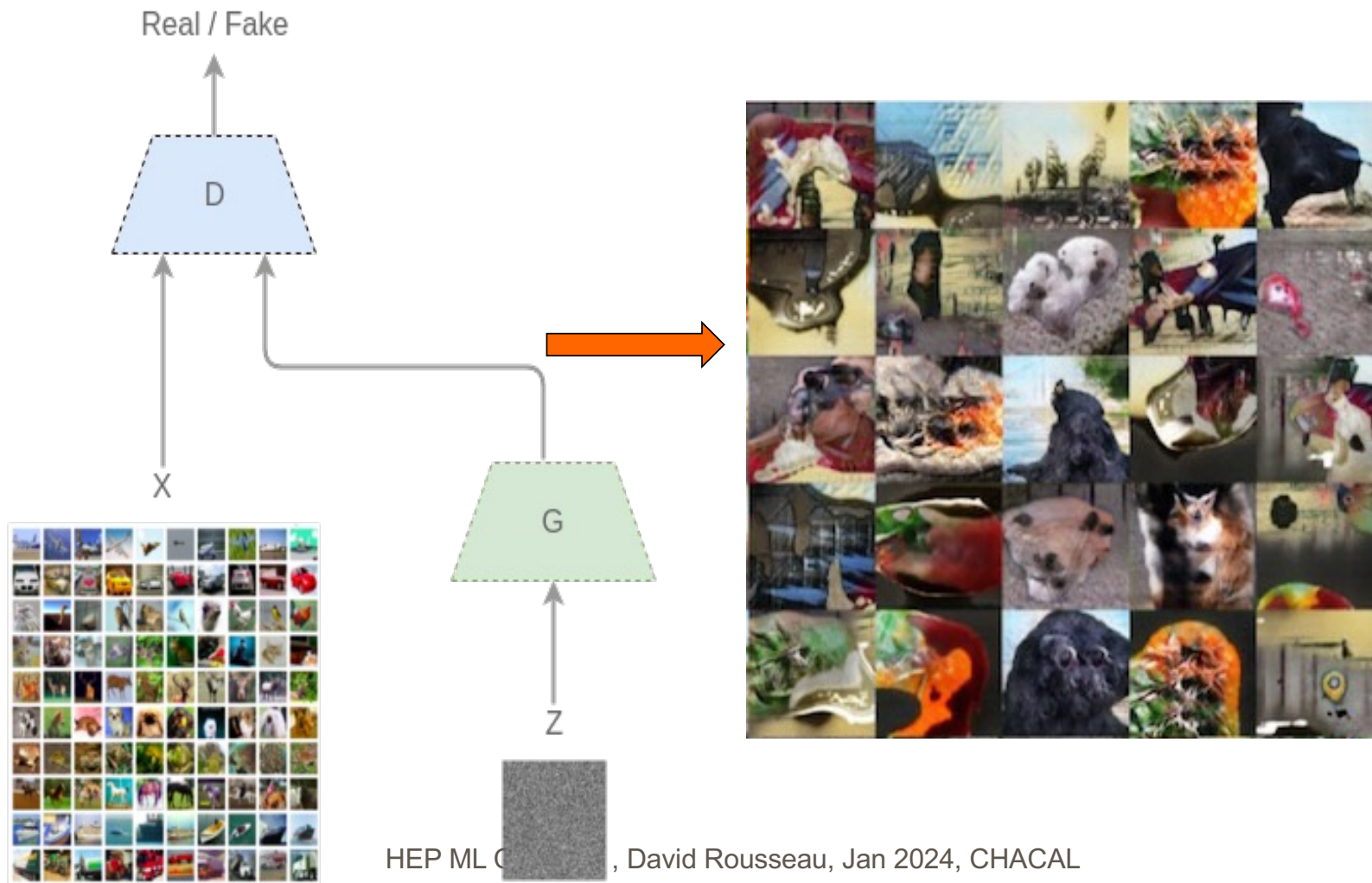
- Extension of previous HiggsML challenge from 2014 (which was a classification problem for Higgs decaying to Tau leptons in an ATLAS simulation based on momenta of decay particles and derived quantities)
- Larger dataset, and include systematic variations
 - Tau Energy Scale, Jet Energy Scale, MET, backgrounds or simulators - TBD)
 - Systematics implemented both using in Delphes detector simulation and post-hoc scripts to compare
- Participants submit methods that go beyond classification: predict signal strength (μ) and an uncertainty interval
- Metric/score on precision and accuracy
- To run on Codabench summer 2024 as a NeurIPS competition



Generative Models



Generative Adversarial Network



Condition GAN



Text to Image

2016

this small bird has a pink breast and crown, and black primaries and secondaries.



the flower has petals that are bright pinkish purple with white stigma



this magnificent fellow is almost all black with a red crest, and white cheek patch.



this white and yellow flower have thin white petals and a round yellow stamen



2024



Here is the image of the small bird with a pink breast and crown, and black primaries and secondaries.

📄 👍 🗨️ 🔄
Actually using diffusion model rather than GAN

Jan 2024



<https://thispersondoesnotexist.com>

<https://www.whichfaceisreal.com/>

PLAY

ABOUT

METHODS

LEARN

PRESS

CONTACT

BOOK

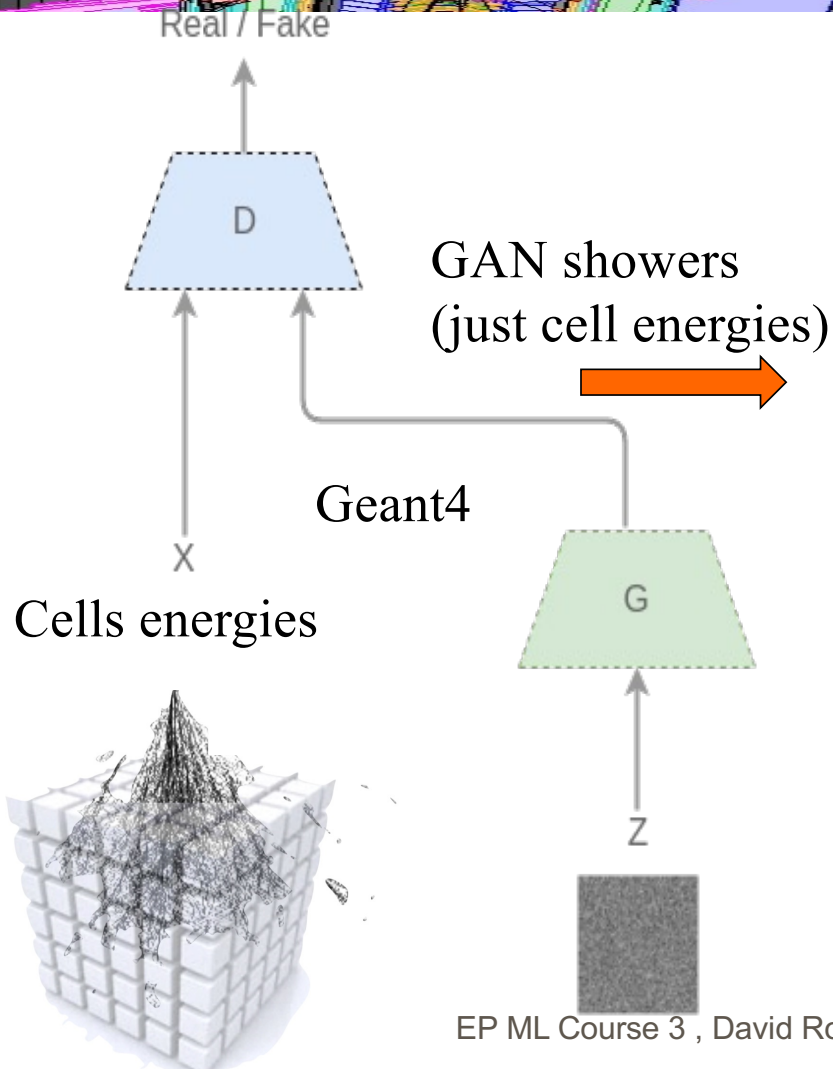
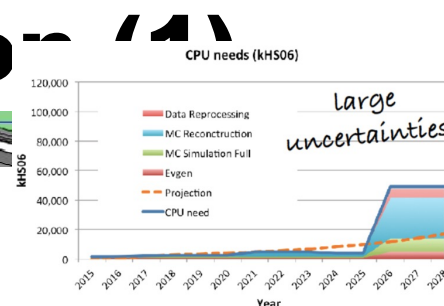
CALLING BS

Click on the person who is real.



HE

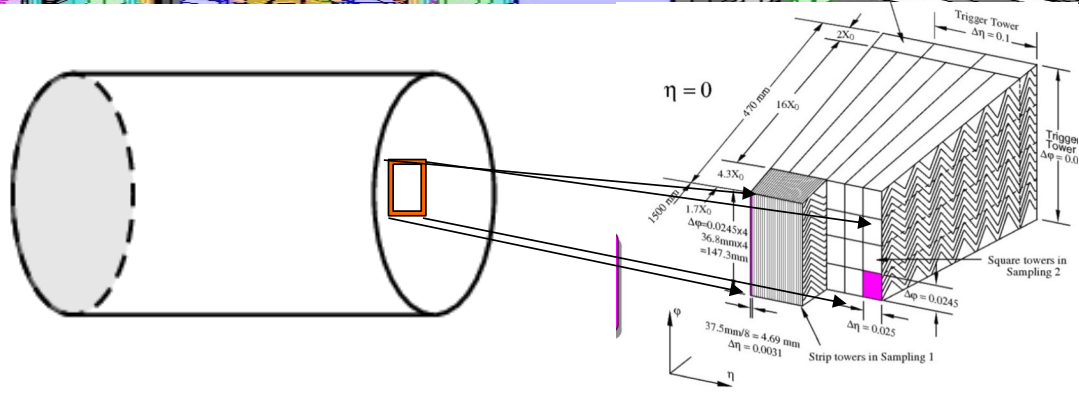
GAN for simulation



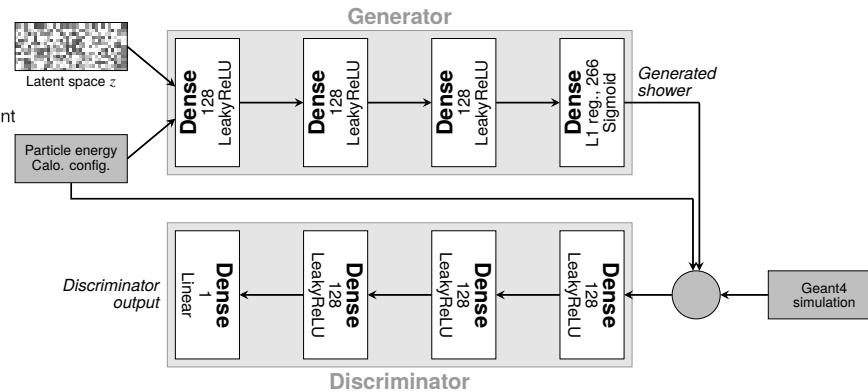
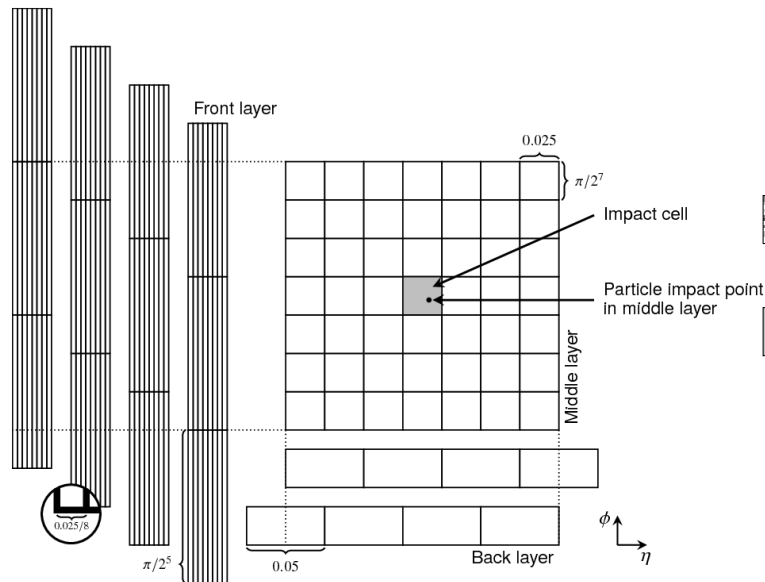
- ❑ Half of LHC grid computers (~300.000 cores) are crunching Geant4 simulation 24/24 365/365
- ❑ ...while LHC experiments are collecting more and more events
- ❑ →reducing CPU consumption of simulation is very important
- ❑ Imagine training a GAN on single particle showers of all types and energies
- ❑ Then when an event is simulated it would ask for GAN showers on request (superfast by 3-4 order of magnitude)
- ❑ Would replace current fast simulation, frozen shower libraries....
- ❑ If/when it works, would require large GPU clusters

ATLAS calo simulation

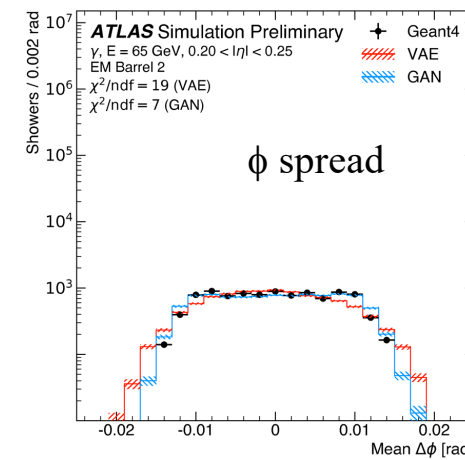
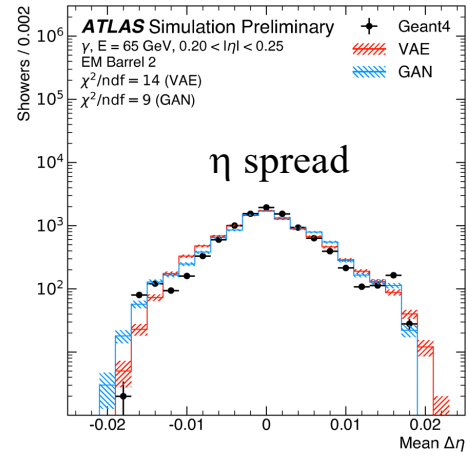
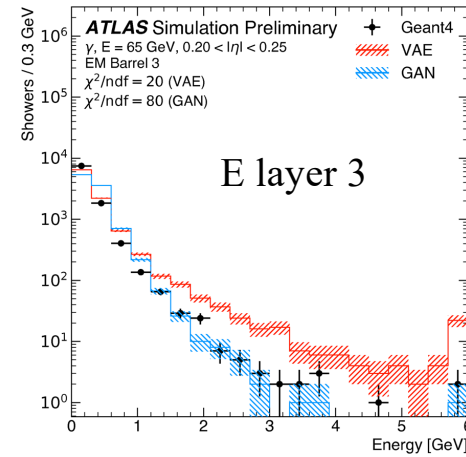
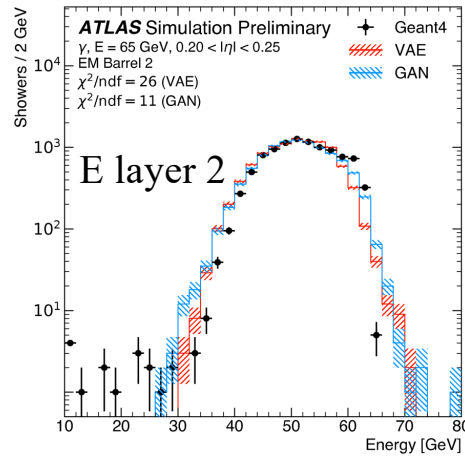
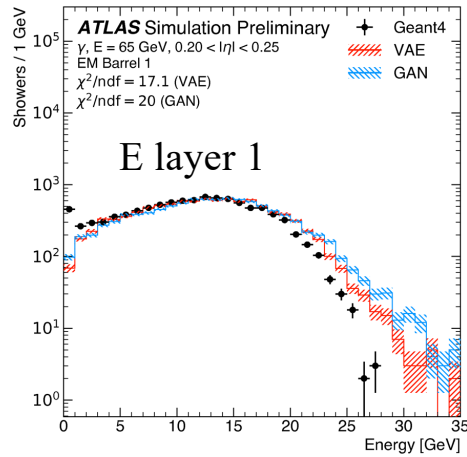
arXiv:2210.06204



+ η, ϕ translation
 177000 cells \rightarrow 266 cells

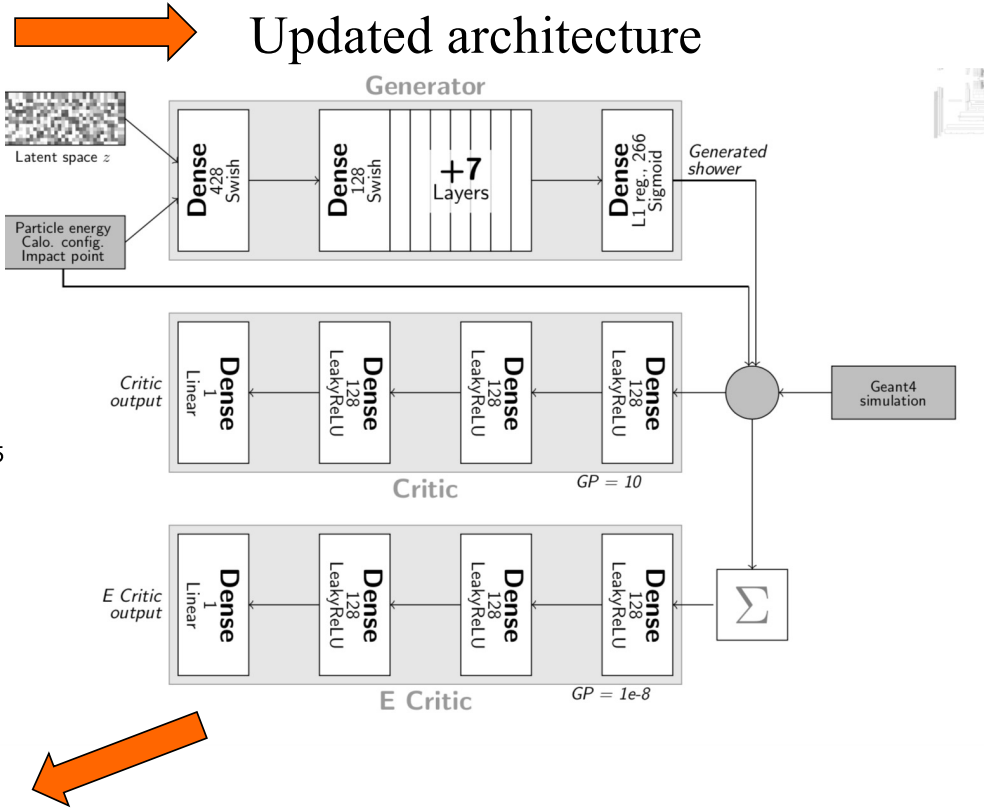
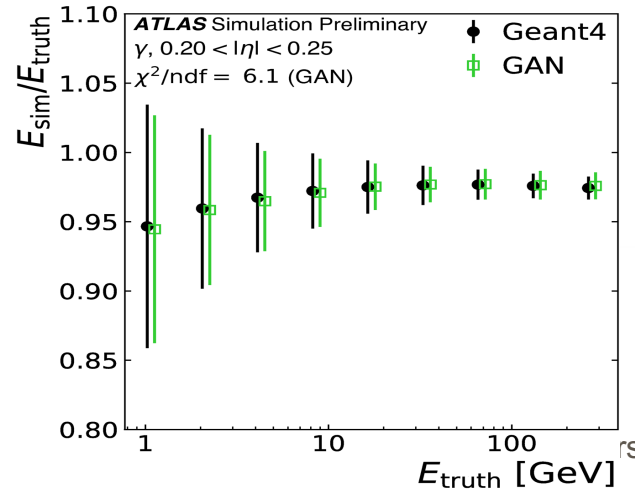
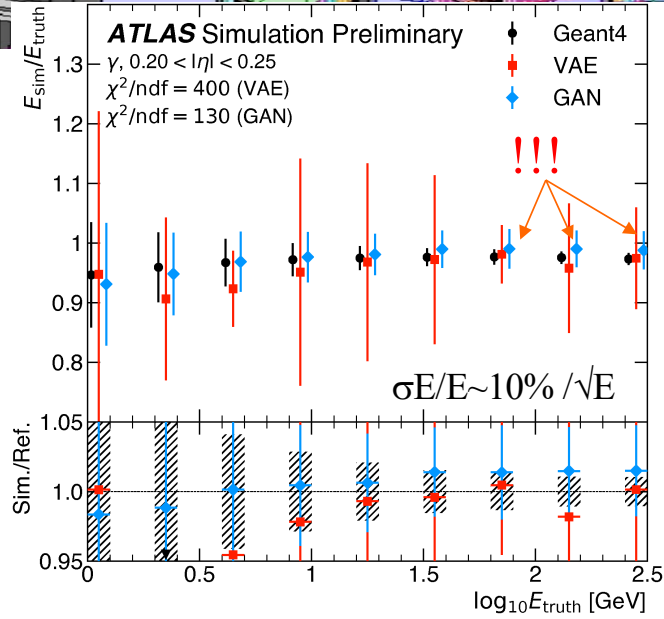


Results



Speed: < 1 ms compared
 to 10s
 Not accurate enough yet

Simulation of energy resolution



Another GAN application

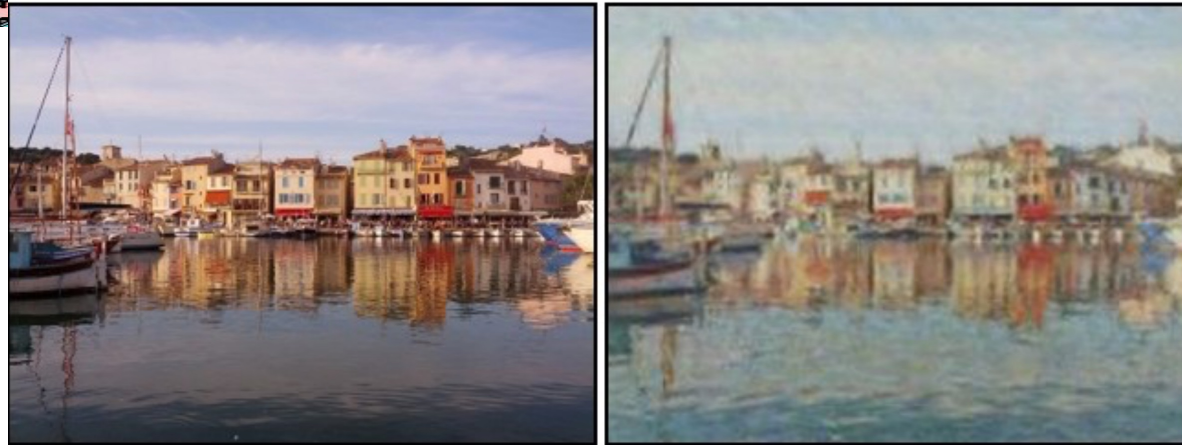


photo → Monet



Monet → photo

HEP ML Course 3, David Rousseau, Jan 2024, CHACAL



CycleGAN

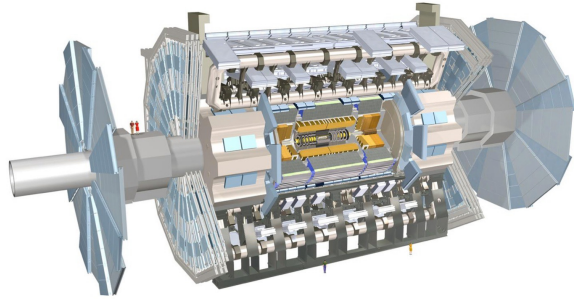




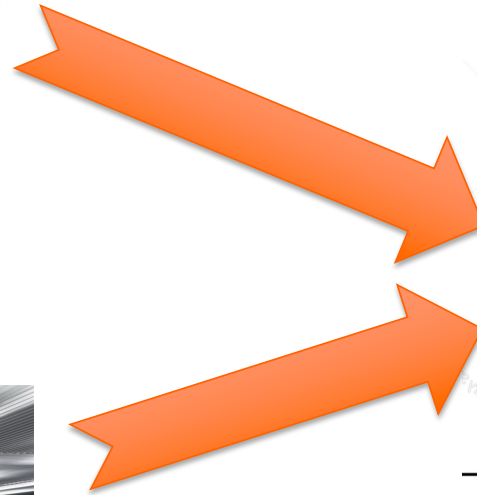
CycleGAN



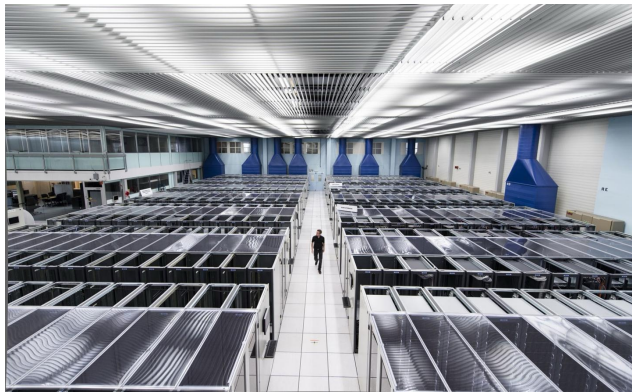
Application



Real data



Simulation



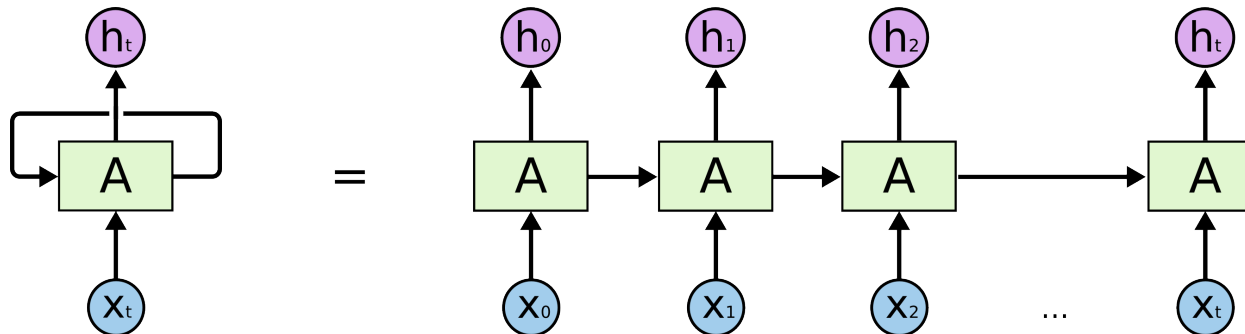
Style Transfer to correct Simulation
→ reduce systematic uncertainties

Recurrent Neural Networks



Recurrent Neural Network

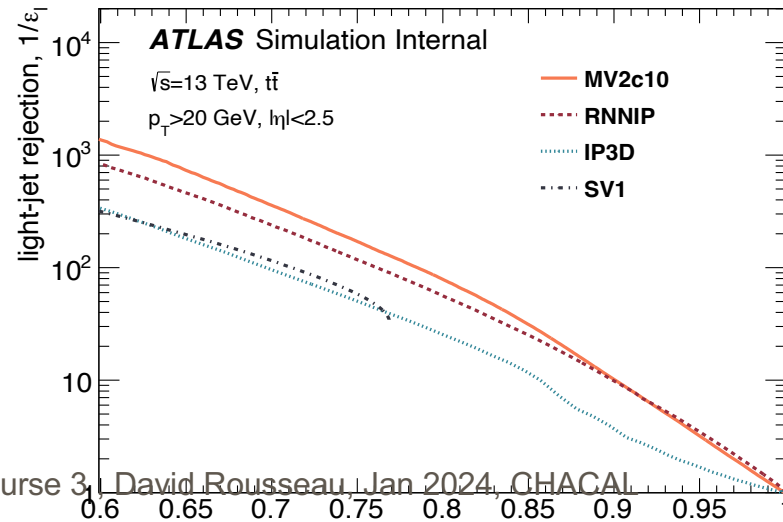
- ❑ BDT and usual NN expect a fix number of input. What to do when the number of inputs is not fixed like list of particles ?
- ❑ Can be used to treat list of particles define pt_1 pt_2 $pt_3...$ ϕ_1 ϕ_2 $\phi_3...$ OK for 1 and 2 particles but does not scale
- ❑ Recurrent Neural Networks have seen outstanding performance for processing sequence data, typically to treat text.
 - Take data at several "time-steps", and use previous time-step information in processing next time-steps data feed into RNN
- ❑ Take list of particles, with well defined order (e.g. decreasing energy or distance to origin) and feed the RNN



RNN for jet tagging

ATL-PHYS-PUB-2017-003

- ❑ Used in physics for identification of jets from b-quarks from the list of particles associated to the jet
 - Basic track information like d_0 , z_0 , pt-Fraction of jet, ...
 - Physics inspired ordering by d_0 -significance
- ❑ RNN outperforms other IP algorithms
 - No explicit vertexing, still excellent performance
 - First combinations with other algorithms in progress



Transformers



Reinforcement learning



Apprentissage par renforcement



Alpha Zero : a appris tout seul contre lui-même en quelques jours à jouer (séparément) au Go, aux échecs ou aux échecs japonais, et bat tout le monde, homme ou machine

RL (1)

States & Actions

Current state of the system



Actions of the player



RL (2)

Rewards & value

Immediate reward

total Value (including the future)

Reward r^t = score at time t as a result of the action



RL (3)

Rewards & value

Immediate reward

total Value (including the future)

Reward r^t = score at time t as a result of the action

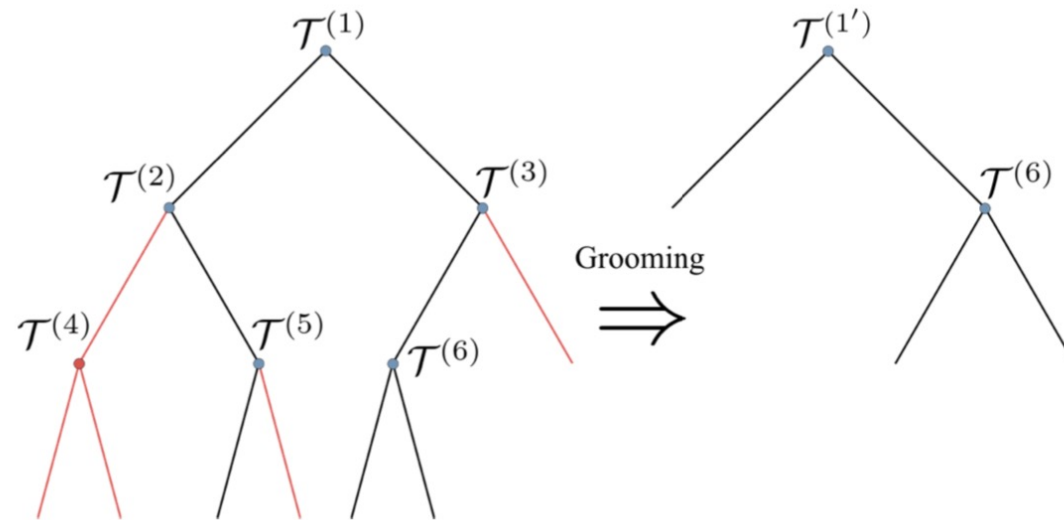
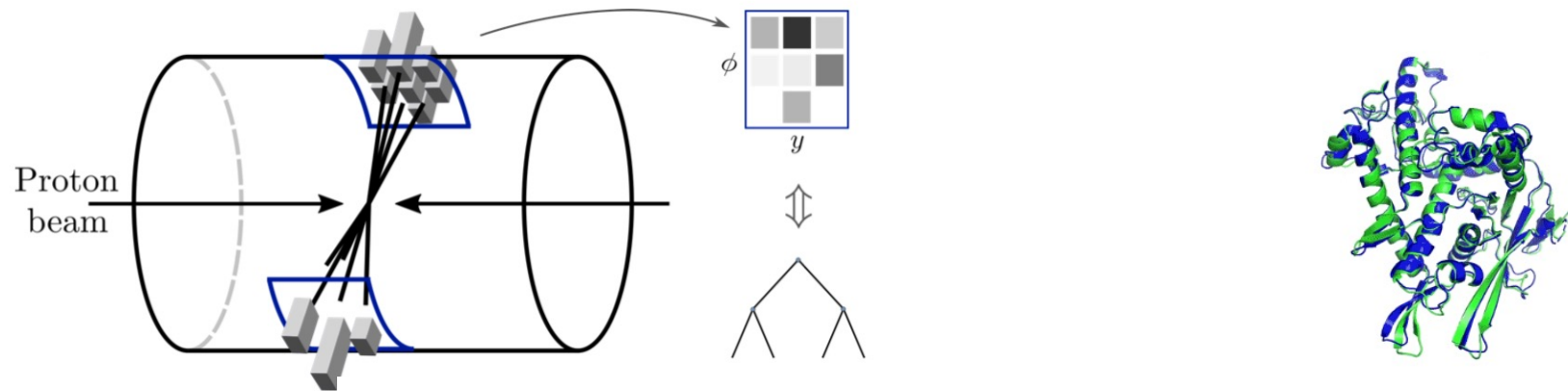
the “Value” of the position
includes future rewards

$$V(\{a_t, s_t\}) = \sum_{\tau=t}^{\infty} \gamma^{\tau} r(\tau)$$



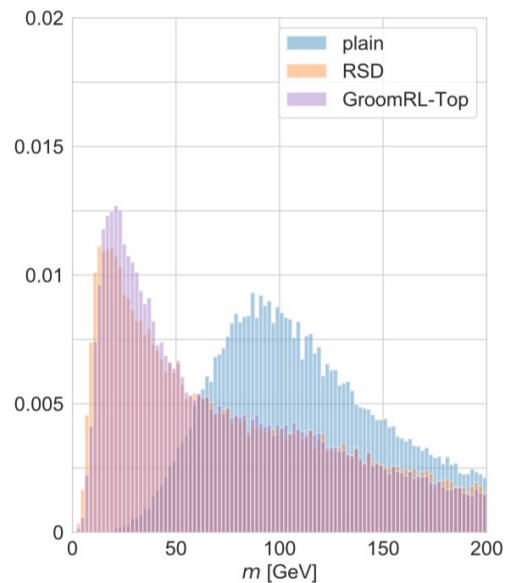
Application of jet grooming

Carrazza, Dreyer 2019

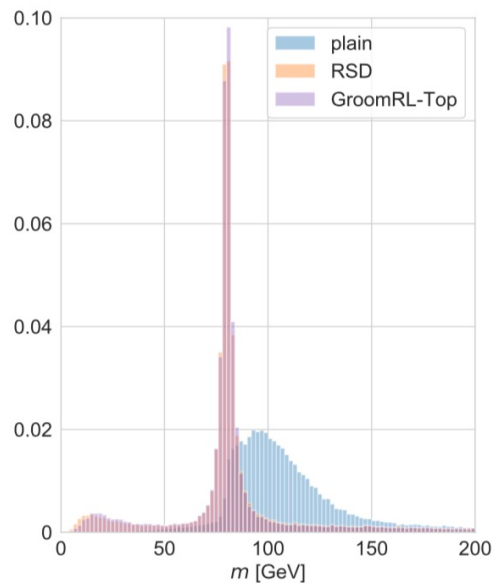


Results

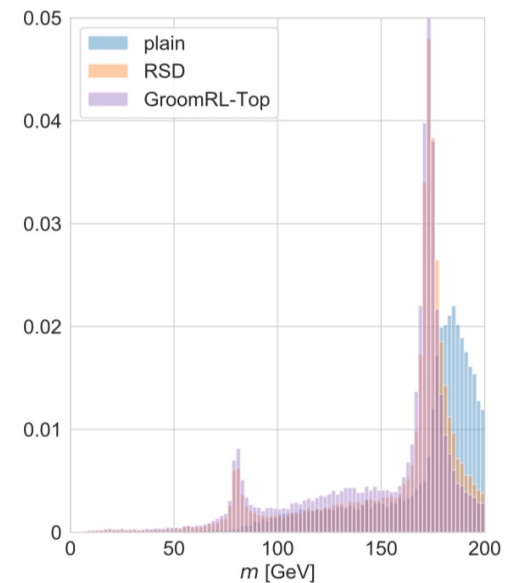
- Train a « policy » (keep a branch or not) according to « reward » : best jet mass
- Impact : more precise jet reconstruction
- (no follow up to the best of my knowledge)



(a) QCD



(b) W



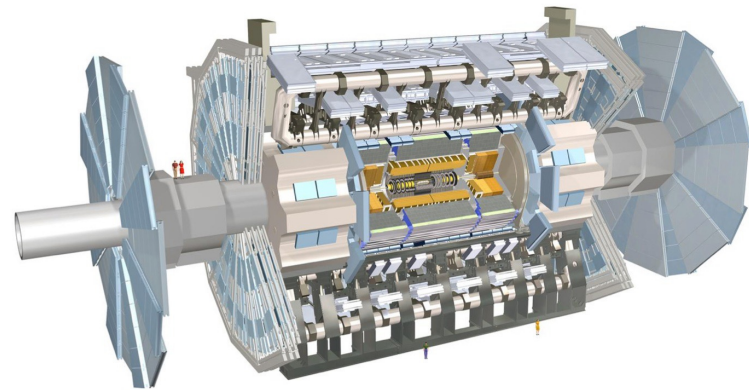
(c) top

Possible RL application : experiment design

- Experiment design
- Data collection
- Success/failure ?



→ could be used to optimise trigger menu, which can be easily virtualised



MODE collaboration

Example of a bad ML idea: SuperTML

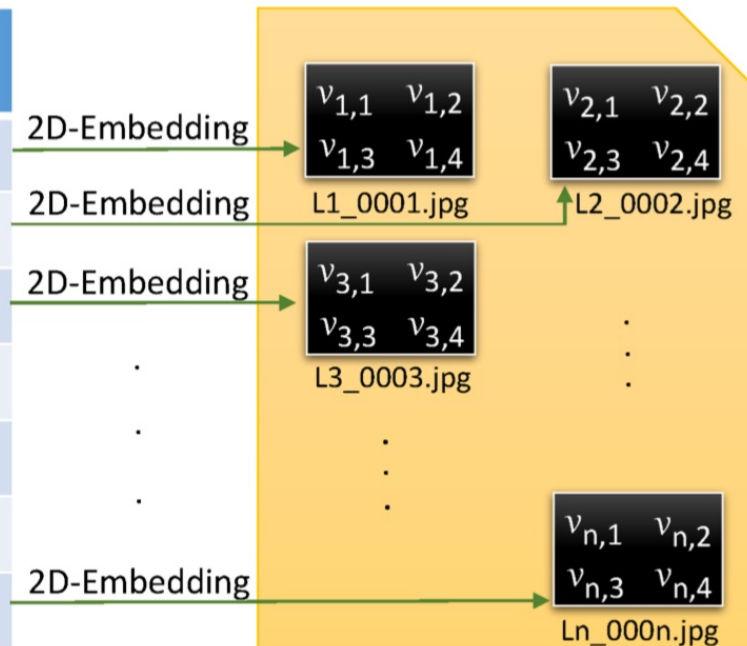


SuperTML : principle

Tabular Data

Features Samples	F1	F2	F3	F4	Label
Sample_1	$v_{1,1}$	$v_{1,2}$	$v_{1,3}$	$v_{1,4}$	L1
Sample_2	$v_{2,1}$	$v_{2,2}$	$v_{2,3}$	$v_{2,4}$	L2
Sample_3	$v_{3,1}$	$v_{3,2}$	$v_{3,3}$	$v_{3,4}$	L3
.
.
.
Sample_n	$v_{n,1}$	$v_{n,2}$	$v_{n,3}$	$v_{n,4}$	Ln

Image Folder



...then analyse image with fine tuned a pre-trained CNN (on cats and dogs)

Super TML on HiggsML dataset



103.706	2.618
50.754	0.602
51.627	310.982
33.558	-1.251
2.047	617.204
2.606	0.945
495.832	-0.693
134.523	1.986
1.973	686.787
93.708	2.214
-1.12	426.564
0.0	0.191
2.238	39.975
4.582	500.637
3	-1.552

(a) SuperTML_EF background event example.

160.937	68.768
103.235	3.473
42.014	44.704
2.078	2.039
0.725	164.546
0.879	1.414
48.146	0.501
36.918	0.103
-1.916	125.157
-999.0	-3.011
1.158	46.226
-999.0	-999.0
-999.0	-999.0
1	46.226
	-999.0

(b) SuperTML_VF signal event example.

Figure 4. Examples of SuperTML images for Higgs Boson .

SuperTML results



Table 3. Comparison of AMS score on Higgs Boson. The first two rows are winners in the Higgs Boson Challenge.

Methods	AMS
DNN by Gabor Meli	3.806
XGBoost	3.761
SuperTML_EF(224x224)	3.979
SuperTML_VF (224x224)	3.838

No Way!
Overtraining ?

...unreproducible....

ChatGPT, a brilliant but messy intern



Caveat : Most experiments from June 2023

GPT in one slide

- ❑ provided by (not at all Open) OpenAI, funded mainly by Microsoft (hence connection to Bing and github Copilot)
- ❑ Generative Pre-Trained Transformer : trained to guess the next word
- ❑ learned from a well curated dataset (details unknown). Certainly not “the whole web”, but wikipedia, arXiv, books etc. Multilingual. Frozen in Sep 2021. → now (Jan 2024) April 2023
- ❑ Fine Tuned to specific task (now GPTs)
- ❑ RLHF : Reinforcement Learning from Human Feedback to provide structured answers
- ❑ safeguards against hate speech etc... Kids can use it
- ❑ GPT3.5 $\sim 10^{11}$ parameters. Memory 10.000 “tokens” (entity a bit smaller than a word)
- ❑ GPT4 $\sim 10^{12}$ parameters. Memory 25.000 tokens

ChatGPT and Bing in one slide

- ❑ ChatGPT : <https://chat.openai.com/chat> register
 - GPT3.5 free, GPT4 for 24\$/month.
 - Does not look up the web, knowledge frozen in Sep 21 (but evidence earlier arXiv are missing)
 - No sources. Reference and citation often wrong
 - On the prompt one can reference a paper giving its title/authors (NOT the http! Wrong paper)
 - Tends to be very verbose
 - History of all different chats are kept and can be continued
 - <https://platform.openai.com/ai-text-classifier> supposed to detect code written with GPT
- ❑ Bing : simply tag @bing in skype or within MS Edge
 - uses GPT4
 - my feeling: looks up on the web and interpret the first few hits
 - →sources are often quoted and correct
 - →not as thorough as ChatGPT
 - newtopic ⇒starts new conversation and forgets the current one (which is still visible skype mode, however Bing answers are deleted after ~10 days)
 - only text input but it is possible to indicate a web page with a pdf (arXiv or on google drive)
 - for more details [this set of rules](#) seems to be followed still, and help understands Bing's behavior

ATLAS study



- ❑ Call for contributions within ATLAS collaboration (1992 created, international, ~3000 physicists including ~200 with AI experience)
- ❑ → ~20 stories from experienced users of ChatGPT/Bing
- ❑ Beyond anecdotes, can ChatGPT/Bing be useful to physicist' daily work ?
 - → a clear yes !
 - However one always has to stay on top of it
 - However would benefit a specialised fine-tune

Public knowledge database



- ❑ ChatGPT more thorough than Bing, but sources missing are a real problem (plus ChatGPT's training DN frozen in Sep 2021) → now (Jan 2024) essentially fixed
- ❑ Very good for textbook knowledge on A or B
- ❑ Not so good to treat A and B, especially if cutting-edge (i.e. science and AI)

Private Knowledge database

← → ↻ <https://indico.cern.ch/event/1306120/> !!!!!!!

CHACAL 2024

15-27 janv. 2024
Fuseau horaire Africa/Johannesburg

- Accueil
- Ordre du jour
- Liste des contributions
- Inscription
- List of students

CHACAL 2024:
and Applications CNRS
15-27 Jan 2024, University of Witwater

☐ → big potential if « fine tuning » to internal knowledge (public papers and corresponding internal discussions, internal documents, mailing lists, meeting slides... since ~2000)

☐ (many startups are proposing companies to build a ChatGPT for their internal documentation)

☐ Indico : >1 million CERN meetings with slides attached since 2000

☐ → incredible, unique, dataset about science as it is built

« Can Normalizing Flows be used at the LHC? »

ChatGPT

Yes, normalizing flows can be used at the Large Hadron Collider (LHC) for various tasks, such as data analysis and simulation. Normalizing flows are a type of generative model that can be used to learn the probability distribution of high-dimensional data. In the context of particle physics, normalizing flows can be used to model the probability distribution of particle collision events.

One application of normalizing flows in particle physics is in the simulation of particle collisions. Simulating particle collisions is essential for predicting the outcomes of experiments at the LHC. However, simulating particle collisions is a computationally expensive task that can require significant computing resources. Normalizing flows can be used to speed up the simulation of particle collisions by learning the probability distribution of the particles produced in a collision and generating new samples from that distribution. This can reduce the computational resources required to simulate particle collisions, which can be especially useful for rare events.

Normalizing flows can also be used for data analysis at the LHC. For example, they can be used to learn the probability distribution of signal events and background events in a particle physics experiment. This can be useful for identifying signal events, which are events that are of particular interest to researchers, and distinguishing them from background events, which are events that are produced by other processes and can obscure the signal events.

In summary, normalizing flows can be a useful tool for various tasks in particle physics, including simulation and data analysis at the LHC.

Bing

Normalizing Flows are generative models which produce tractable distributions where both sampling and density evaluation can be efficient and exact 1. They are used for distribution learning 2.

I found a research paper titled "Normalizing Flows for LHC Theory" which discusses the use of normalizing flows in LHC theory 3. However, I couldn't find any other information about the use of normalizing flows at LHC.

Is there anything else I can help you with?

1. <https://arxiv.org/abs/1908.09257>
2. <https://arxiv.org/abs/1505.05770>
- 3.

← best primary reference

https://www.researchgate.net/publication/368556193_Normalizing_Flow_s_for_LHC_Theory

- Bing has tried to make sense from the first few hits
- **ChatGPT more thorough but absence of source a clear handicap**
- ...however one should ask follow-up questions (A then B):

How can normalizing flows be used to learn probability distribution exactly ?

(...)

thanks, please provide as a full latex document

document

Let's suppose we have a dataset \mathcal{D} of N samples, each of which is a high-dimensional vector $\mathbf{x} \in \mathbb{R}^D$. The goal of a normalizing flow model is to learn a transformation $f_\theta : \mathcal{Z} \rightarrow \mathcal{X}$, parameterized by θ , that maps a simple base distribution, such as a Gaussian distribution, to the target distribution.

Specifically, we want to learn the transformation such that the probability density of the transformed samples $p_{\mathcal{X}}(f_\theta(\mathbf{z}))$ matches the true probability density of the dataset $p_{\mathcal{D}}(\mathbf{x})$.

The transformation f_θ is constructed by composing a sequence of invertible mappings f_i :

$$f_\theta = f_K \circ f_{K-1} \circ \dots \circ f_2 \circ f_1$$

Each f_i is an invertible function that takes as input a vector \mathbf{u}_i and outputs a vector \mathbf{v}_i :

$$\mathbf{v}_i = f_i(\mathbf{u}_i)$$

The inverse function of f_i is denoted by g_i , such that $g_i(\mathbf{v}_i) = \mathbf{u}_i$. The log-likelihood of a sample \mathbf{x} can be computed using the change of variables formula:

$$\log p_{\mathcal{X}}(\mathbf{x}) = \log p_{\mathcal{Z}}(\mathbf{u}) - \sum_{i=1}^K \log \left| \det \frac{\partial f_i}{\partial \mathbf{u}_i} \right|$$

where $\mathbf{u} = g_K \circ g_{K-1} \circ \dots \circ g_2 \circ g_1(\mathbf{x})$ is the inverse transformation of \mathbf{x} , and $\frac{\partial f_i}{\partial \mathbf{u}_i}$ is the Jacobian matrix of the transformation f_i with respect to \mathbf{u}_i . During training, the parameters θ are learned by maximizing the log-likelihood of the dataset \mathcal{D} with respect to θ :

$$\theta^* = \arg \max_{\theta} \sum_{\mathbf{x} \in \mathcal{D}} \log p_{\mathcal{X}}(\mathbf{x}; \theta)$$

This optimization can be performed using gradient-based methods, such as stochastic gradient descent.

Jan 2024: chatGPT much better



Vous

Can Normalizing Flows be used at the LHC? Please provide references



ChatGPT

Normalizing Flows can indeed be used in the context of the Large Hadron Collider (LHC) for various tasks such as data analysis, simulation, anomaly detection

Here are some references that detail their application:

- Data Analysis and Event Reconstruction:** Normalizing Flows are used for high-precision simulations and evaluating uncertainties in high-energy physics. This includes applications like event reconstruction [1].
- Fast Simulation:** There is research on using generative models and Normalizing Flows, for detector response modeling at the LHC, leading to faster and efficient simulations [2].
- Anomaly Detection:** Studies discuss the improvement of anomaly detection strategies at the LHC using variational autoencoders and Normalizing Flows [3] [4].
- Noise Reduction and Data Compression:** Normalizing Flows are used for high-dimensional detector simulations, which include tasks like noise reduction in complex datasets [5].
- Uncertainty Quantification:** The application of Normalizing Flows for uncertainty quantification in high-energy physics simulations is also discussed in the literature [6].

NO!

Article

Full-text available

Normalizing Flows for LHC Theory

February 2023 · *Journal of Physics Conference Series* 2438(1):012004

DOI: [10.1088/1742-6596/2438/1/012004](https://doi.org/10.1088/1742-6596/2438/1/012004)

License · [CC BY 3.0](https://creativecommons.org/licenses/by/3.0/)

Anja Butter

Normalizing Flows for High-Dimensional Detector Simulations

Florian Ernst^{1,2}, Luigi Favaro¹, Claudius Krause^{1,3}, Tilman Plehn¹, and David Shih⁴

Improving Variational Autoencoders for New Physics Detection at the LHC With Normalizing Flows

2022 | Jawahar, Pratik; Aarrestad, Thea; Chernyavskaya, Nadezda; Pierini, Maurizio; Wozniak, Kinga A; Ngadiuba, Jennifer; Duarte, Javier; Tsan, Steven

Published Web Location

<https://doi.org/10.3389/fdata.2022.803685>

Generative Machine Learning for Detector Response Modeling with a Conditional Normalizing Flow

Allison Xu

Shuo Han

Xiangyang Ju

Haichen Wang

Jan 2024, CHACAL

Scientific reading

- mediocre: « What are the important points of this paper? » → all points correct, but
 - Some do not come from the paper (but from general knowledge)
 - « important » is too vague
- However specific questions can be correctly answered:
« why is technique A used » ?
 - → in general 80/20 correct vs BS

Scientific writing

- ❑ Very good at suggesting a paper structure/substructure
- ❑ Useless for the actual writing (we're writing about cutting edge science not available elsewhere...)
- ❑ Not bad at summarising (80/20)
- ❑ Good at outreach: I am in high school, please explain the Higgs boson to me



what would be a good analogy for the Higgs boson if I am 18 ?



The Higgs boson is a particle that helps give other particles mass, so here's an analogy that might help you understand it better:

Imagine a school hallway where students are moving around, but some of them are able to move faster than others. The faster students are like particles that have less mass, while the slower students are like particles with more mass. Now imagine that there's a teacher in the hallway, and every time a student passes by the teacher, the student gets slowed down a bit. This teacher is like the Higgs field, which particles interact with to acquire mass. The more a particle interacts with the Higgs field, the more massive it becomes.

And just like how some students might have more trouble getting past the teacher and thus end up moving slower, some particles have more trouble interacting with the Higgs field and thus end up being more massive.

Code

- ❑ Very good at writing simple code snippets
 - → 80% correct
- ❑ Can also explain/debug code in unfamiliar language →
- ❑ Good at using well known libraries
- ❑ Not as good for specific code (3 M lines of C++/Python in ATLAS) → would need a dedicated « fine tuning »

I get an error when I try:

```
$ conda env create -f cpu_environment.yml python=3.9 -n exatrk-cpu
Solving environment: failed

CondaValueError: Malformed version string '~': invalid character(s).
```

It looks like there might be an error in the `cpu_environment.yml` file that you are trying to use to create your conda environment. The error message you are seeing suggests that there might be an issue with the version number specified in the environment file.

To resolve this issue, you could try removing the `python=3.9` argument from your command and see if that works. Here is an updated version of the command:

HEPML_HandsOn_NN.ipynb

Fichier Modifier Affichage Insérer Exécution Outils Aide

+ Code + Texte Copier sur Drive Connecter GPU Colab AI

Hands on : introduction to NN on HEP dataset

```
-f cpu_environment.yml -n exatrk-cpu
```

CHACAL