

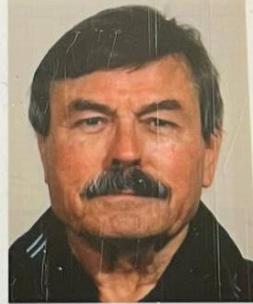


# Некоторые аспекты достоверности результатов ML/DL в экспериментах CERN

[Vassily.Sakharov@cern.ch](mailto:Vassily.Sakharov@cern.ch)

CERN EP-SFT





C-35611

Exp. 08/04/2029

---

Vasily SAKHAROV

---



428745

D

# Искусственный интеллект -

– Машинное обучение -

– Глубокое обучение



# Виды машинного обучения



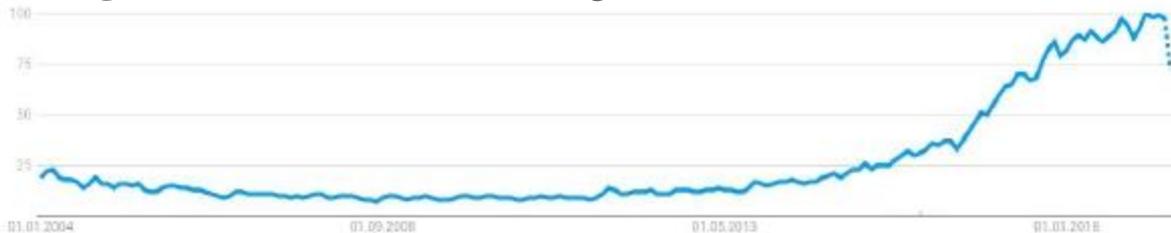
# Машинное обучение в НЕР и прочих сферах

- НЕР и CERN используют методы машинного обучения с 2000 года, традиционно известного как "многомерный анализ "
- Распространенность метода вызвало широкий интерес и появился фирменный термин «машинное обучение»

Google Trends: multivariate analysis



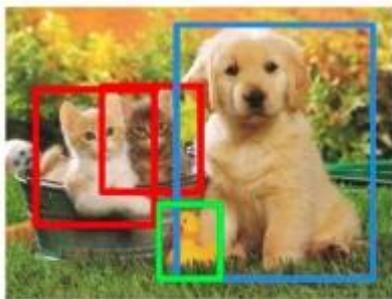
Google Trends: machine learning



# Машинное обучение в НЕР и прочих сферах

## Industry:

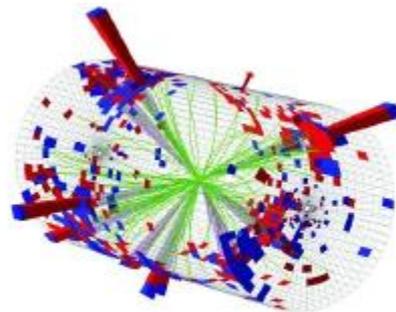
- Обработка естественного языка
- Распознавание речи и почерка
- Обнаружение мошенничества
- Обнаружение спама и вирусов
- Распознавание лиц
- Сегментация и классификация изображений
- Обнаружение объектов...



CAT, DOG, DUCK

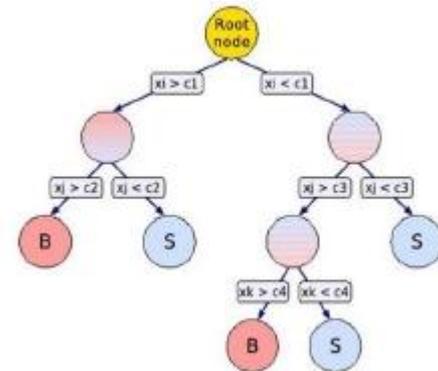
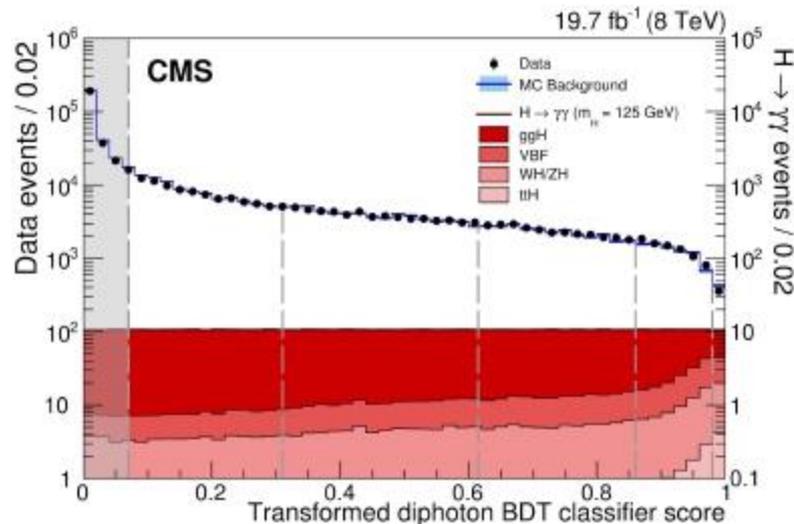
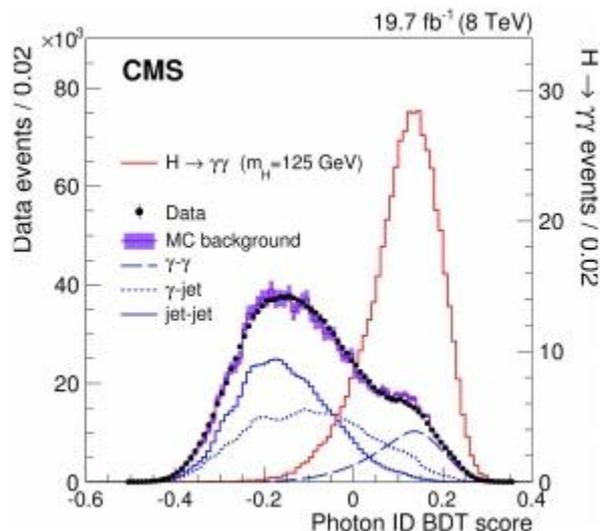
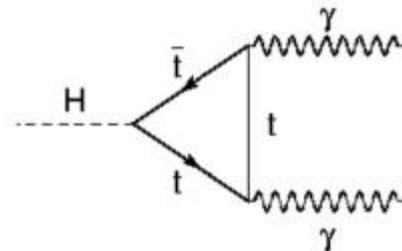
## НЕР:

- Классификация событий
- Реконструкция частиц
- Идентификация частиц
- Калибровка частиц
- Спусковой механизм
- Мониторинг качества данных
- Быстрое моделирование...



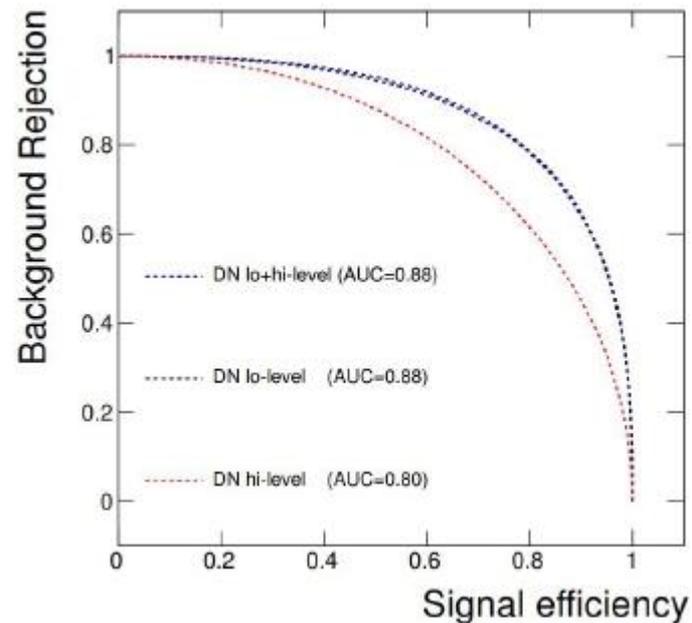
# Классическое машинное обучение в НЕР

- Пример: Наблюдение распада Хиггса в двух фотонах (2012)
- Анализ сильно зависит от доминанты деревьев решений
- Задачи: Идентификация фотонов и классификация событий



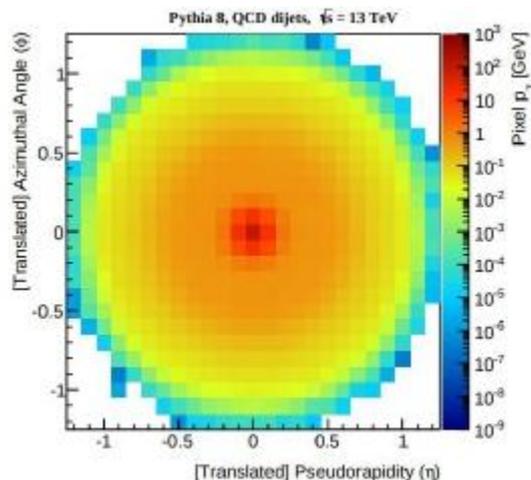
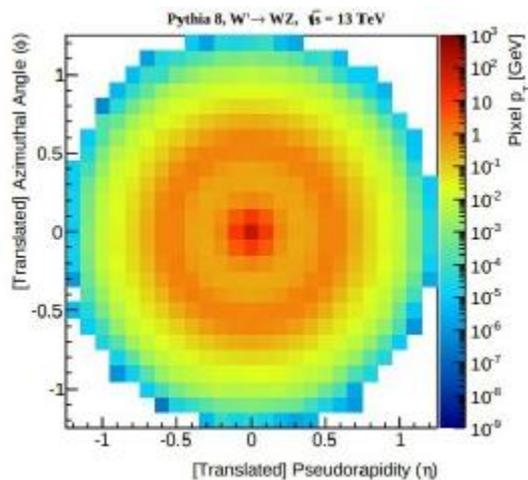
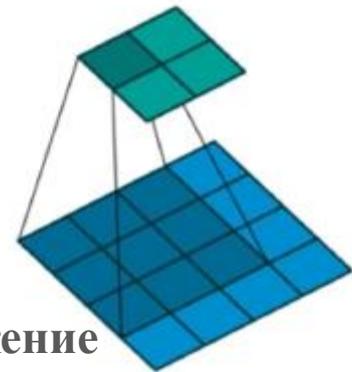
# Глубокое обучение в НЕР

- Традиционный подход: Физик разрабатывает различающие переменные и использует классическое машинное обучение на основе этих параметров и величин
- Тренд: Использование глубоких моделей для обучения и извлечения признаков



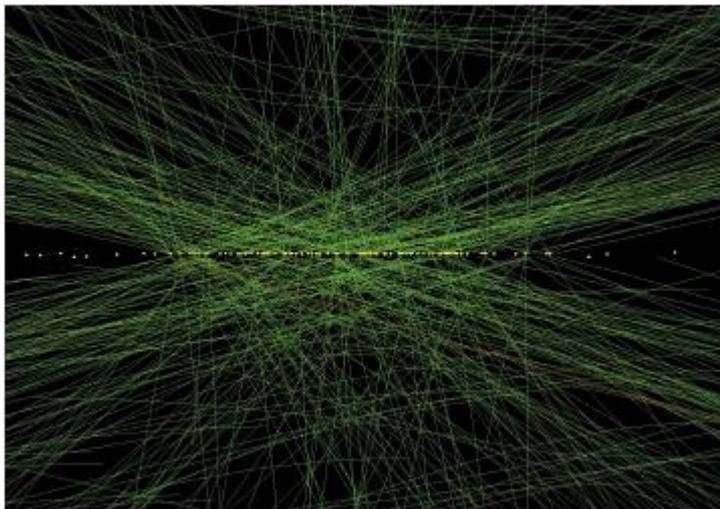
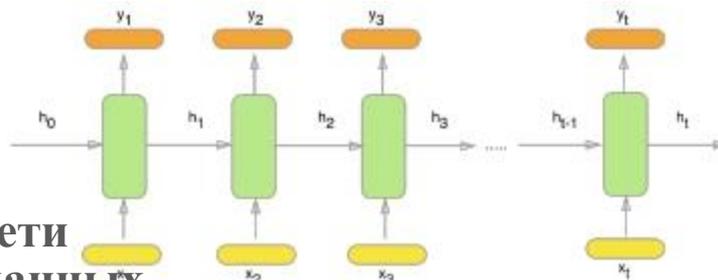
# Сверточные нейронные сети

- Свертка: «Скользящее окно» над изображением
- Традиционное использование : Интенсивно используется в компьютерном зрении
- Использование в НЕР: Детектор, симулирующий 3D-изображение



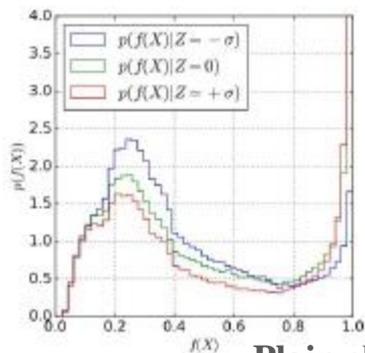
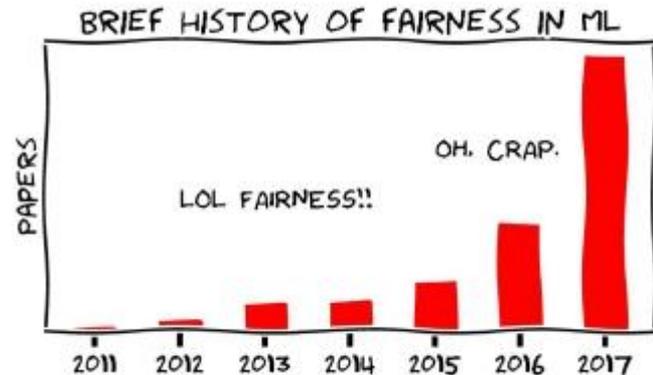
# Рекуррентные нейронные сети

- **Рекуррентность:** Состояние нейронной сети в зависимости от предыдущих входных данных
- **Традиционное использование :** Анализ временных рядов, перевод языка,...
- **Использование в НЕР:** Классификаторы, обрабатывающие входные данные переменной длины, например, треки частиц...

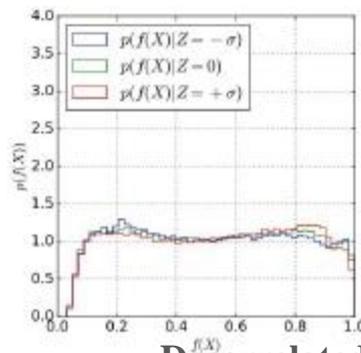
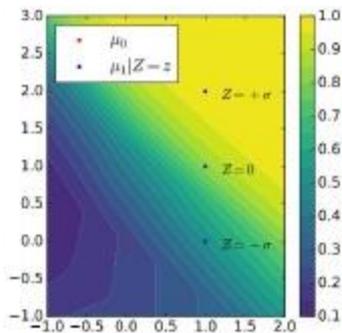


# Состязательные нейронные сети

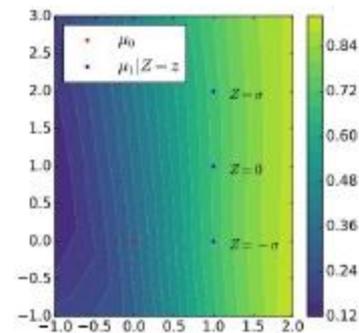
- Построение классификаторов, не зависящих от предопределенных свойств
- Противник: Вторая нейронная сеть используется как бы для подавления зависимости от конкретного свойства
- Традиционное применение : Сделать классификатор независимым от пола, расы и пр. трухи...
- Использование в НЕР: Формирование классификатора, независящего от систематических отклонений/ошибок - **очень важно!**



Plain classifier

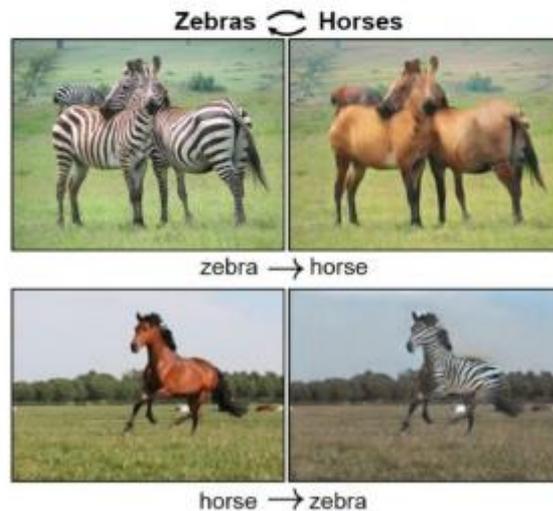
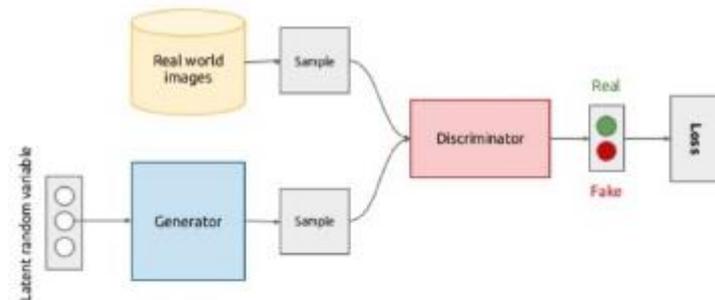
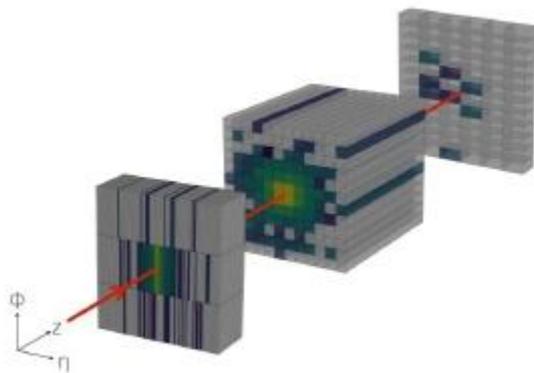


Decorrelated classifier



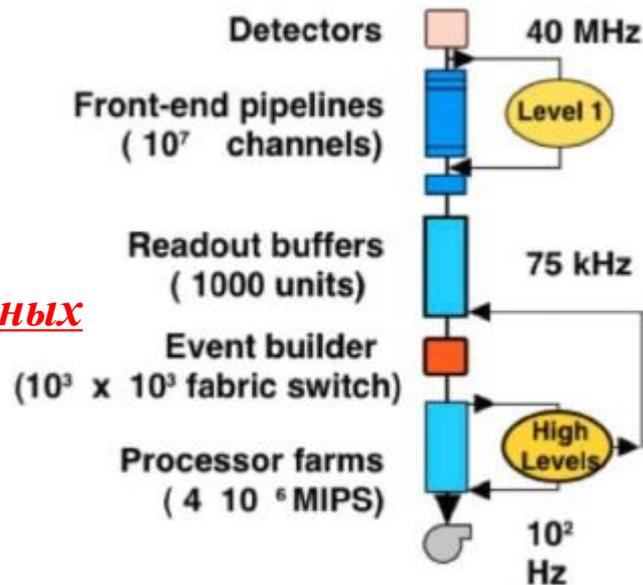
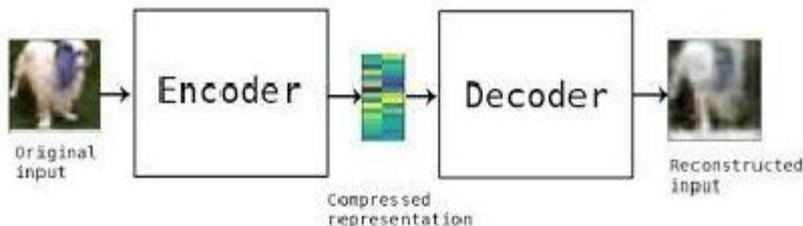
# Генеративные состязательные нейронные сети

- Порождающие модели: создание экземпляров из «белого шума»
- Традиционно: Домен передачи изображений,...
- Использование в НЕР: Генерировать отклик детектора (гораздо быстрее, чем обычные подходы)



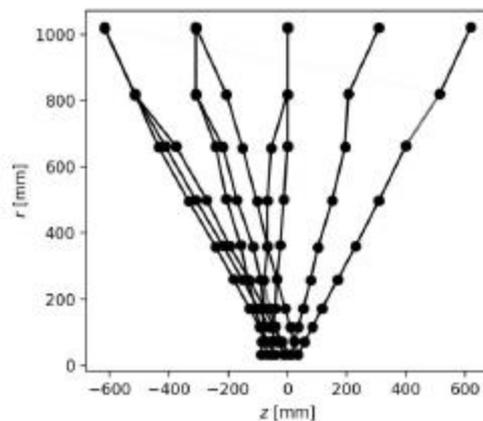
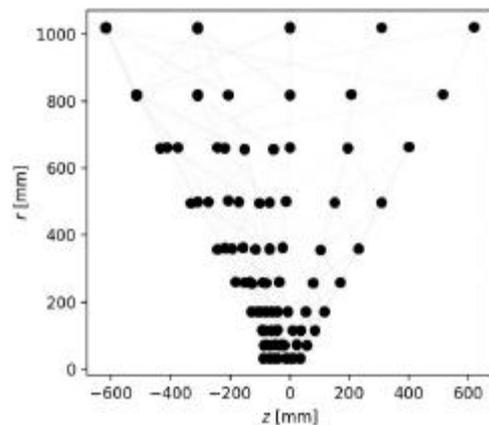
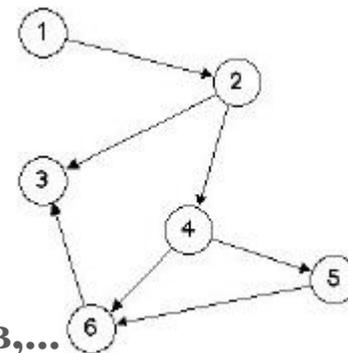
# Глубокий анализатор событий

- Автоэнкодер: Неконтролируемое обучение на разных данных
- Традиционно для: Обнаружения аномалий
- Использование в НЕР: Обнаружение редких событий в *триггере, мониторинг качества данных*



# Графовые нейронные сети

- Нейронные сети, которые могут обрабатывать графы
- Традиционно для: Комбинаторная оптимизация, прогнозирование свойств молекул, классификация узлов,...
- Использование в НЕР: Отслеживание треков, регистрируемых детекторами



# Summary

Детально можно ознакомиться тут

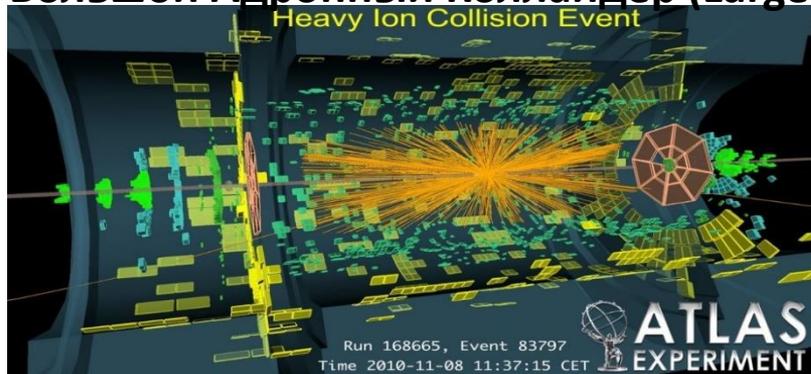
- Machine learning at the energy and intensity frontiers of particle physics  
<https://www.nature.com/articles/s41586-018-0361-2>
- Deep Learning and its Application to LHC Physics  
<https://arxiv.org/abs/1806.11484>

# BackUps

- И еще немного по BIG data in HEP

# Большие данные с детекторов экспериментальной ФВЭ и их обработка

Большой Адронный Коллайдер (Large Hadron Collider - LHC) в ЦЕРНе



Один из 4-х LHC экспериментов - ATLAS



RUN 1  
2012  
Получен  
бозон  
Хиггса!

Схема процесса обработки данных эксперимента

1 петабайт =  $10^{15}$  байт

1 эксабайт =  $10^{18}$  байт

RUN 2 2015-2017 гг = уже  
эксабайты данных в год!

Детекторы LHC-экспериментов выдают данные со скоростью около 1 петабайт/сек

Система интеллектуальных триггеров и фильтров сжимает эти данные в миллионы раз, оставляя на долгое хранение лишь полезную информацию, в итоге БАК выдает для хранения 50 терабайт в секунду, - столько данных за 4 часа, сколько вся сеть Facebook собирает за сутки.

Обработать такой объем данных в ЦЕРНе невозможно, поэтому

1. Создана всемирная интернет-сеть распределенных вычислений (Worldwide LHC Computing Grid - WLCG)
2. Для моделирования и анализа данных разработаны многочисленные пакеты программ, использующих методы машинного обучения

# Итоги и перспективы

- Применение методов машинного обучения было эффективным на всех стадиях развития систем обработки экспериментальных данных ФВЭ, прогрессируя вместе с развитием вычислительных технологий и алгоритмической базы.
- Радикальные проекты последних лет для экспериментов с высокой светимостью (HL-LHC) и NICA, ставят сложную проблему реконструкция треков частиц в плотных средах, для решения которой необходимо развитие новых алгоритмов глубокого трекинга и их распараллеливания на суперкомпьютерах, а также на методах квантового машинного обучения - QML.
- Следует отметить перспективность исследований по применению нейросетевых моделей трансформеров, позволяющих, в частности, эффективно отфильтровывать фейковые измерения и выполнять трекинг на сырых данных, минуя этап с получением хитов.
- В более далекой перспективе следует также уделять внимание методам квантового отжига в приложениях как к глобальному трекингу, так и локальным методам прослеживания, обобщающих алгоритмы фильтра Калмана.
- На волне успеха генеративно-состязательных нейросетей в создании картин и диссертаций следует отметить публикации об их успешном применении для симуляции взаимодействий в экспериментах ФВЭ – **специальные модели LLM в теоретической физике.**

# Big Data Tools and Pipelines for Machine Learning in HEP

# Data Engineering to Enable Effective ML

- From “Hidden Technical Debt in Machine Learning Systems”, D. Sculley et al. (Google), paper at NIPS 2015

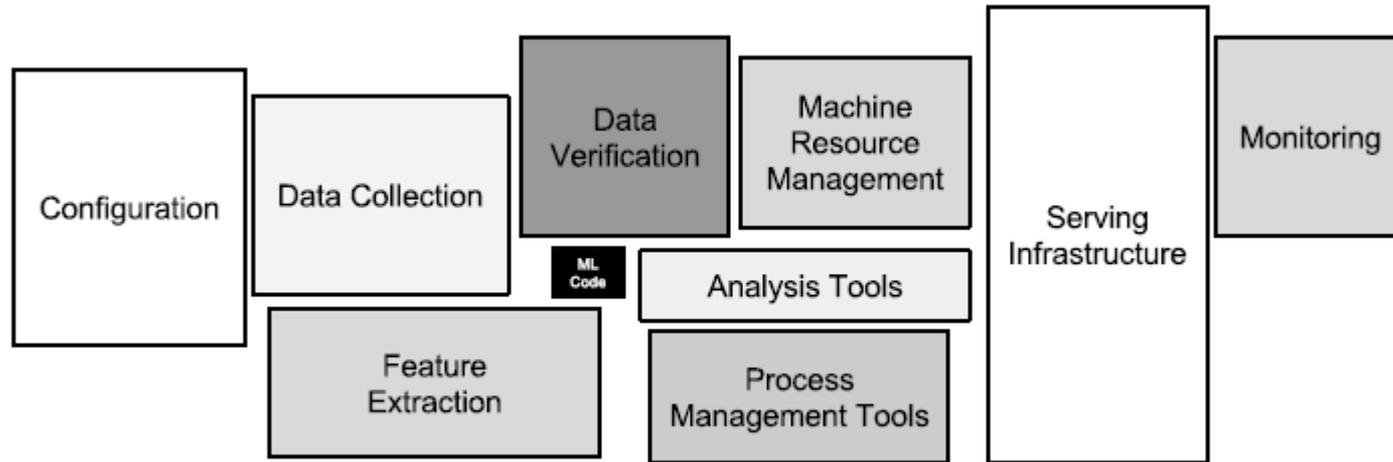
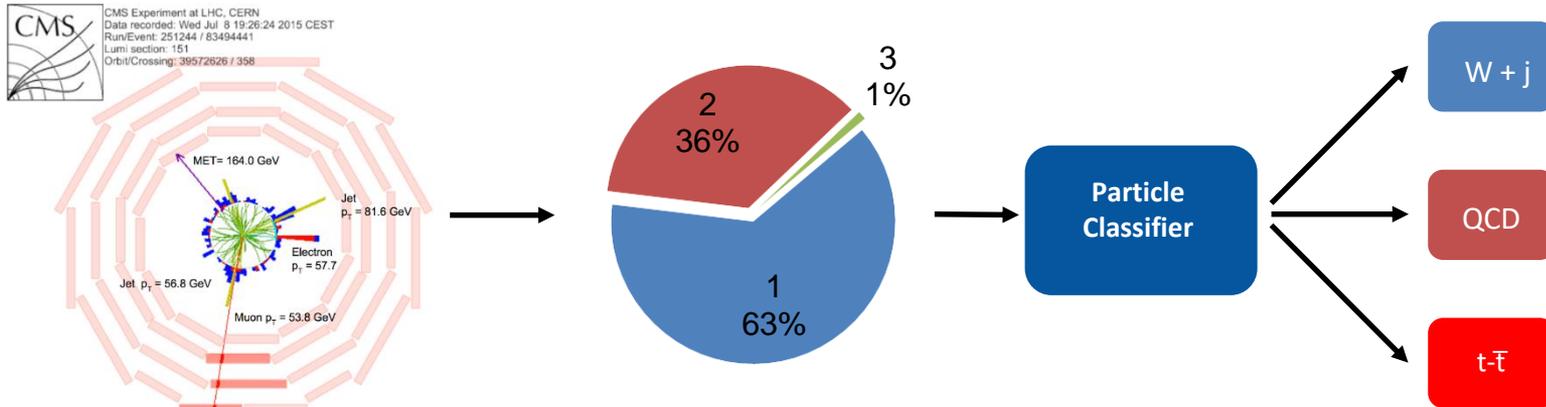


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

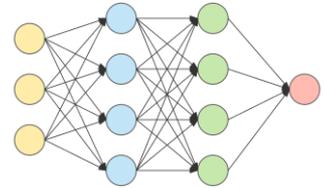
# Particles Classifier Using Neural Networks

- R&D to improve the **quality of filtering systems**
  - **Develop** a “Deep Learning classifier” to be used by the filtering system
  - **Goal**: Identify events of **interest** for physics and reduce false positives
    - False positives have a **cost**, as wasted storage bandwidth and computing
  - “Topology classification with deep learning to improve real-time event selection at the LHC”, Nguyen et al. **Comput.Softw.Big Sci. 3 (2019) no.1, 12**

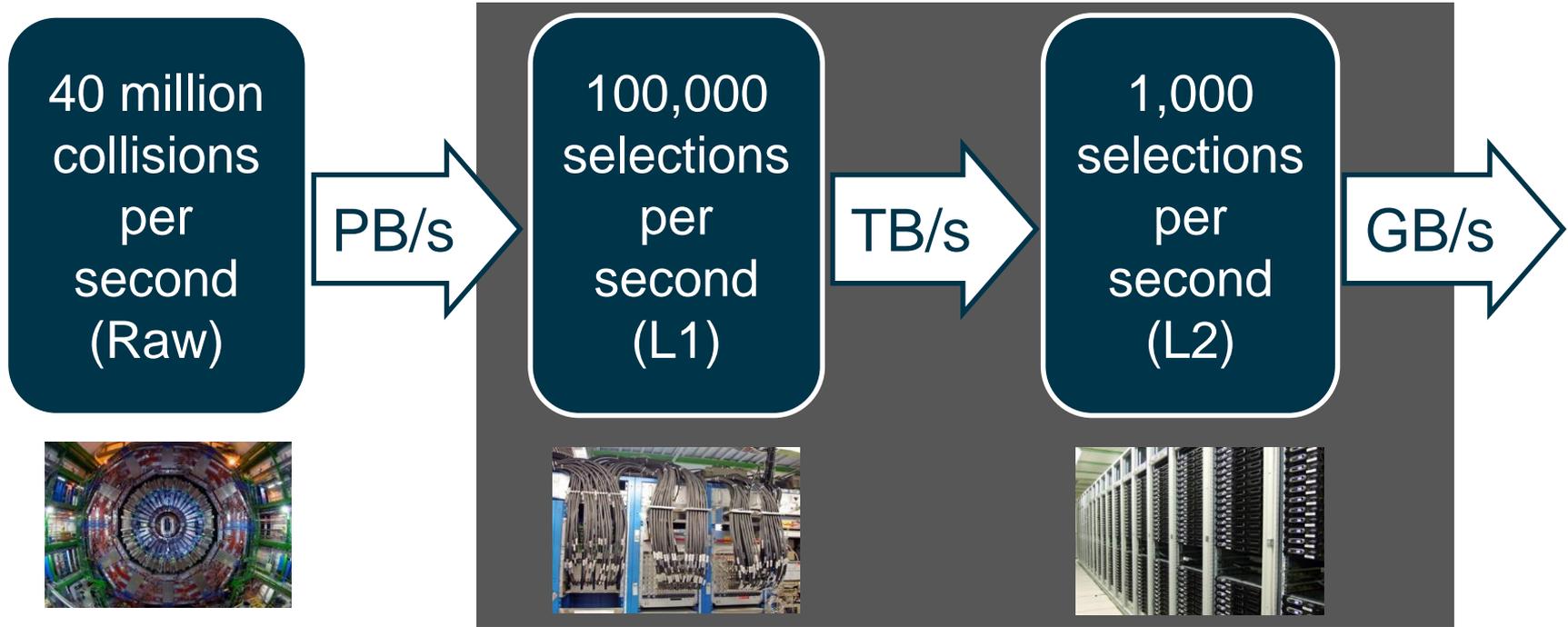


# R&D – Data Pipelines

- Improve the **quality of filtering systems**
  - Reduce false positive rate
  - Complement or replace rule-based algorithms with classifiers based on Deep Learning
- Advanced analytics **at the edge**
  - Avoid wasting resources in offline computing
  - Reduction of operational costs



# Data Flow at LHC Experiments

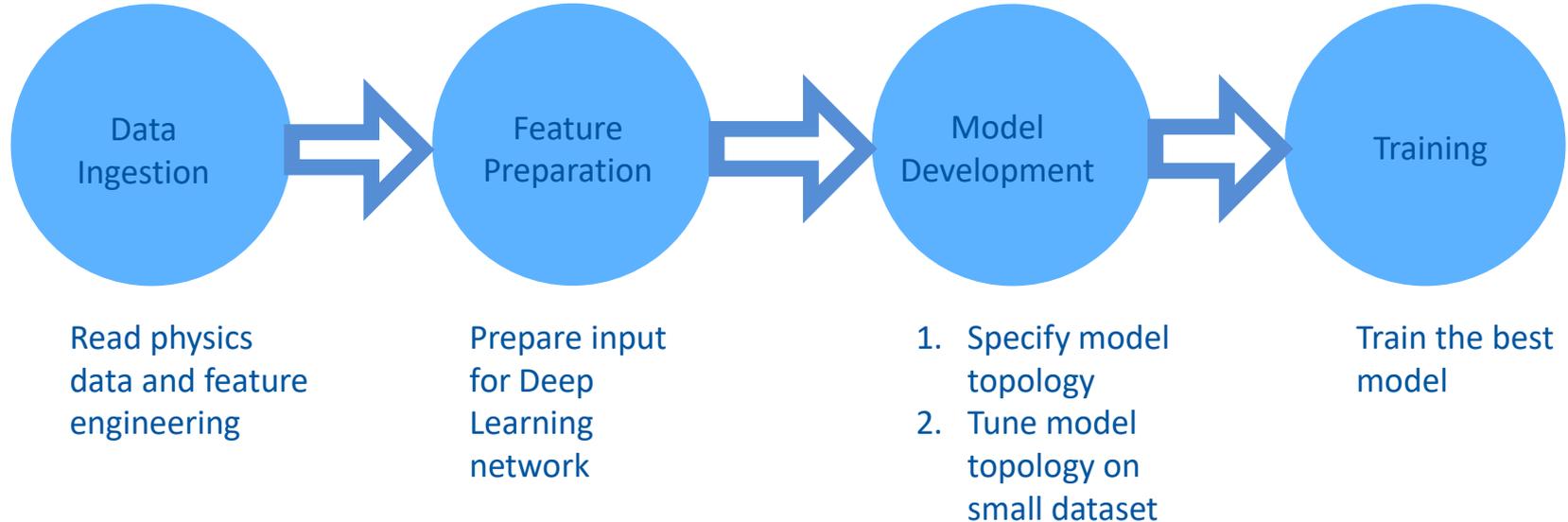


This can generate up to a **petabyte** of raw data per second

Reduced to **GB/s** by filtering in real time

Key is how to select potentially interesting events (**trigger** systems).

# Deep Learning Pipeline for Physics Data



**Technology: the pipeline uses Apache Spark + Analytics Zoo and TensorFlow/Keras. Code on Python Notebooks.**

# Fast Machine Learning at the Edge for HEP Experiments

Sioni Summers

CERN Detector Seminar

8th March 2024

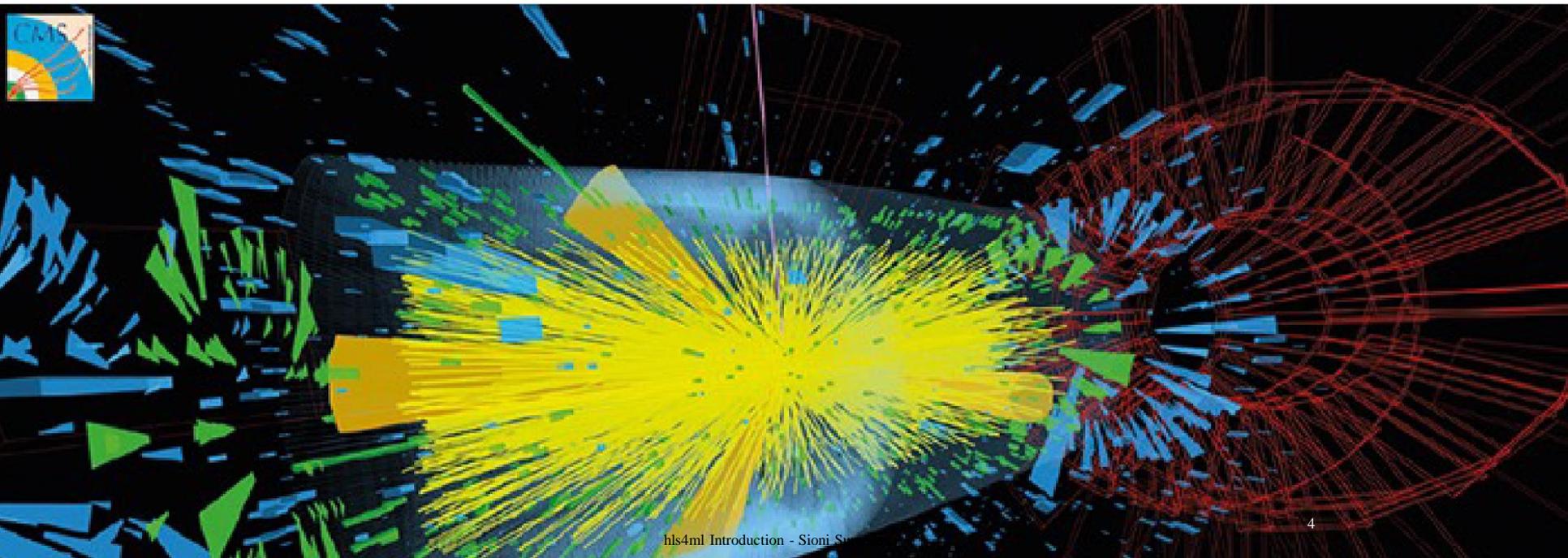


# The challenge: triggering at LHC

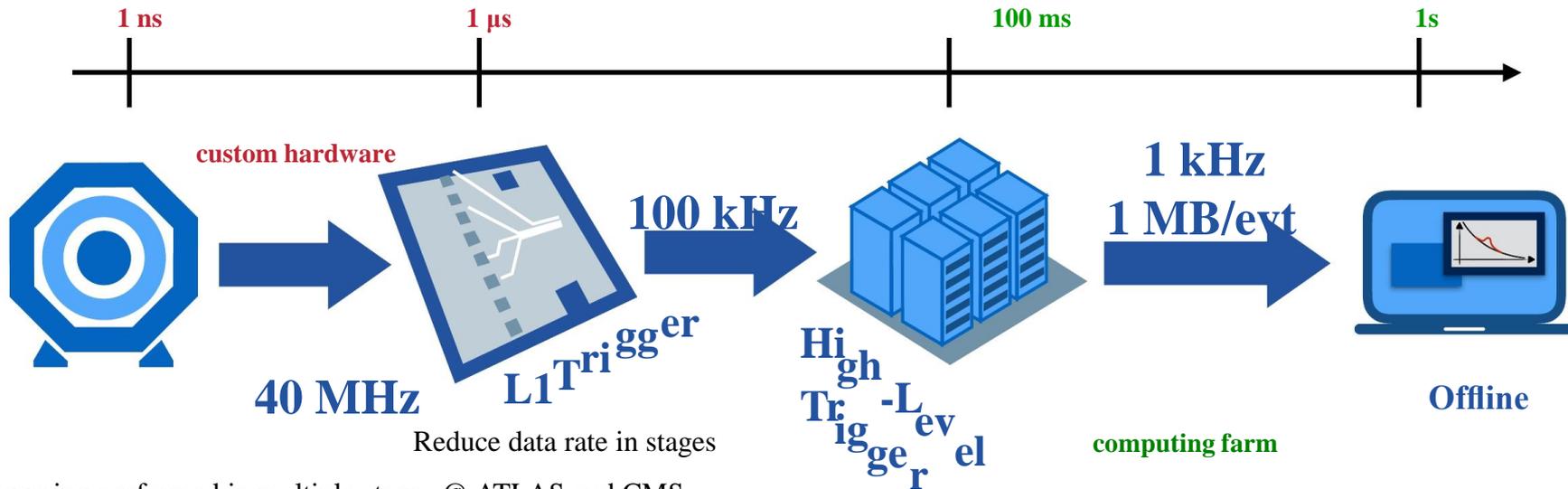
At LHC protons collide at 40 MHz  $\rightarrow$  extreme data rates  $O(100 \text{ Tb/s})$

Most collisions don't produce exciting new particles

**“Triggering” = filtering events to reduce data rates to manageable levels**



# Trigger at LHC



Triggering performed in multiple stages @ ATLAS and CMS

Trigger decision to be made in latency  $O(\mu\text{s})$

Frontends in rad. hard ASICs, processing in

Process 100s Tb/s Trigger (hardware)

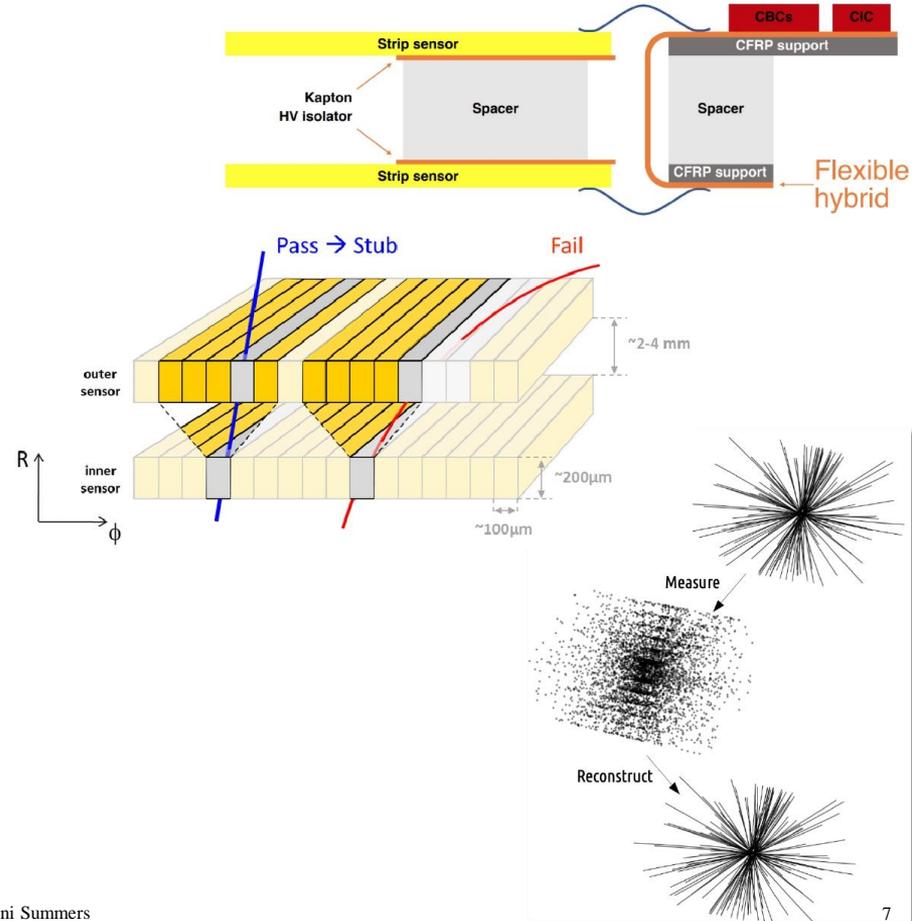
- 99.75% rejected

FPGAs decision in  $\sim 4 \mu\text{s}$

- Trigger farm detailed
- 99% rejected  $O(100 \text{ ms})$
- decision in  $\sim 100\text{s ms}$

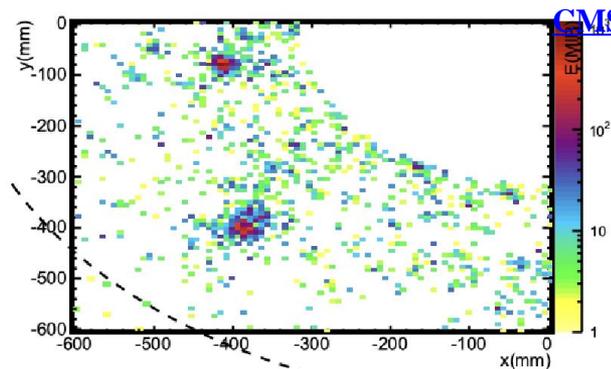
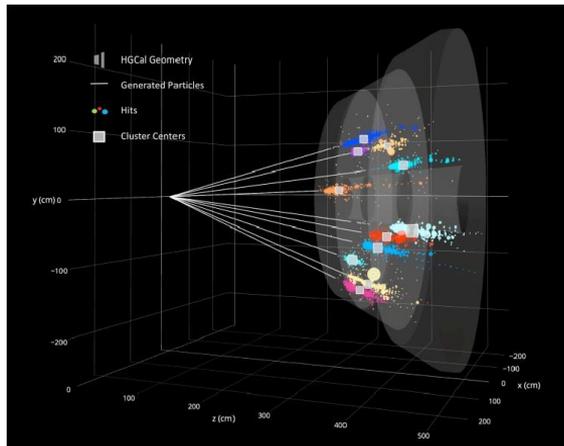
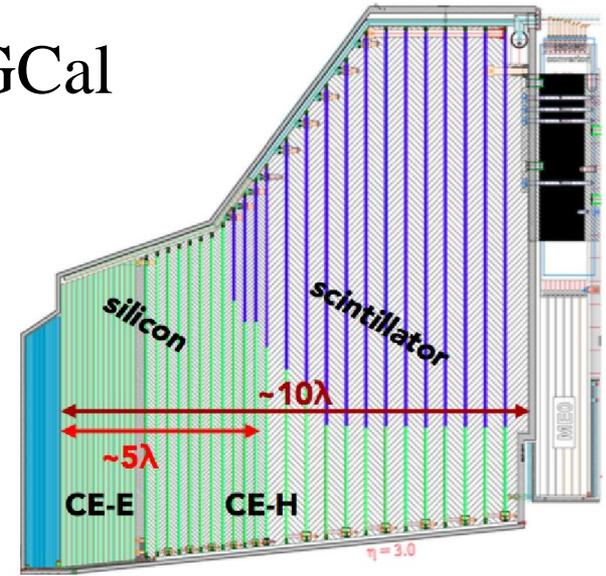
# CMS Detector Upgrade 1: Tracker

- At CMS the Phase 2 Upgrade brings data from new detectors to the trigger for the first time
- New Outer Tracker implements on-detector  $p_T$  cut
- Two silicon strip sensors separated by few mm
  - Far enough to measure bending of charged particle in B field
  - Close enough to be read out on one device
  - “Stubs” passing 2 GeV  $p_{Tcut}$  → Level 1 Trigger
- Level 1 Track Finder system reconstructs tracks from stubs
  - **Around 200 FPGAs, with “classical” tracking algorithms: data organisation, seed building, road following, track fitting**
- Tracks in the Level 1 Trigger essential for 200 PU conditions
  - Reconstruct primary vertex for suppression of 200 PU background
  - Better measurements of properties of electrons, muons, jets
  - 6.4 Tb/s of reconstructed tracks sent downstream



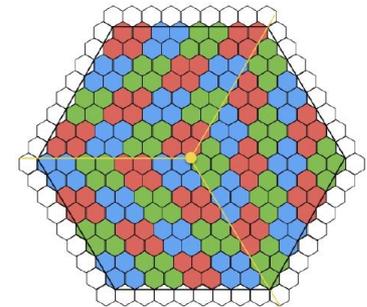
# CMS Detector Upgrade 2: HGCal

- High granularity calorimeter: silicon sampling calorimeter for the endcaps
- 6.5 million channels in 50 layers
  - Very fine transverse and longitudinal segmentation
- **Dedicated ASIC to prepare data for trigger reconstruction - more later**
- **Trigger backend comprises around 200 FPGAs**
- Reconstructing 3D clusters: 4 Tb/s clusters sent downstream



CMS-TDR-019

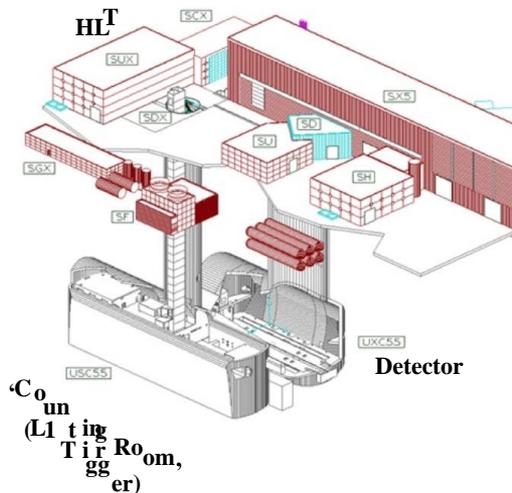
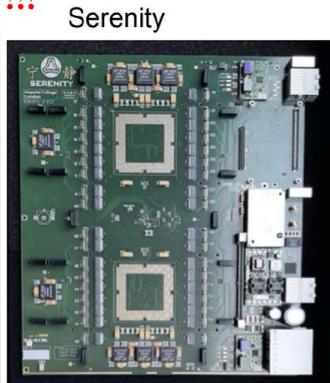
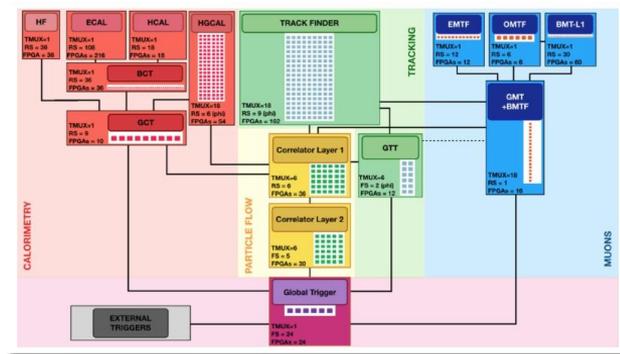
[arXiv:1708.08234](https://arxiv.org/abs/1708.08234)



# CMS Level 1 Trigger

- Phase 2 Upgrade of CMS L1T will have hundreds of boards with FPGAs like those shown below - AMD/Xilinx Ultrascale+ FPGAs
- Data rate of multiple terabits per second into / out of each board on optical fibres
- System organised in layers with normally  $\sim 1\text{-}2 \mu\text{s}$  per step
  - Reducing raw detector data into physics objects (e.g. track finding: hits to tracks)
  - New event every 25 ns, latency for trigger decision for one event 12.5  $\mu\text{s}$

• **Final output is one bit: keep or discard event !!!**



# Machine Learning

- **Build models that learn from data in order to make predictions on new, unseen data**

- “*Models*” can be Neural Networks, Decision Forests, or anything else “trainable”

- “*Training*” is the process of fitting the model parameters that best describe the data

- “*Inference*” is the process of using a fitted model to make new predictions

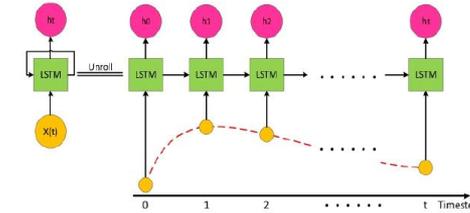
- For Fast ML at experiments we are mostly concerned with making fast inference

- **ML used in HEP since the first ML wave in the 80s, and nowadays extremely prevalent**

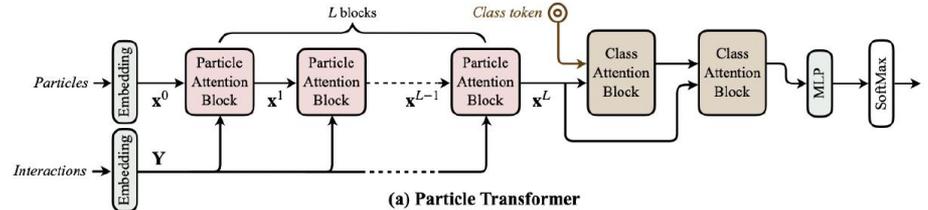
- Different model types for different data representations



Tabular: fully connected or Decision Forest

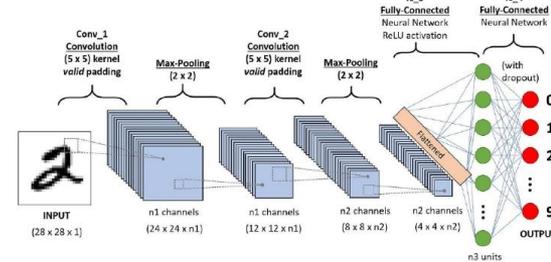


Time series: recurrent

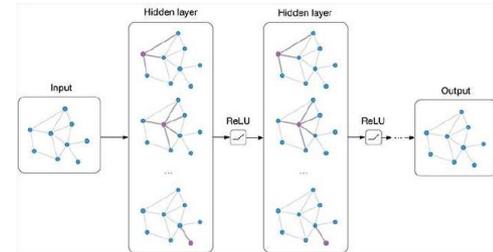


(a) Particle Transformer

Sequence-to-sequence: transformer



Images: convolutional

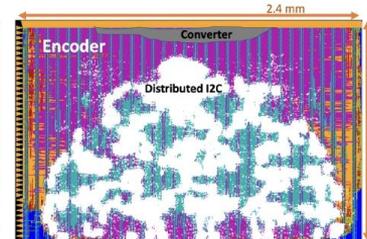
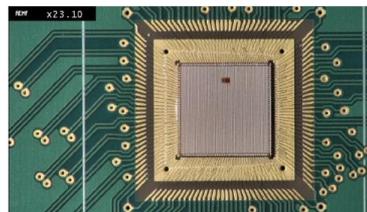
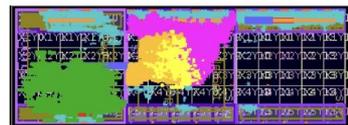
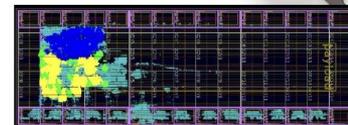
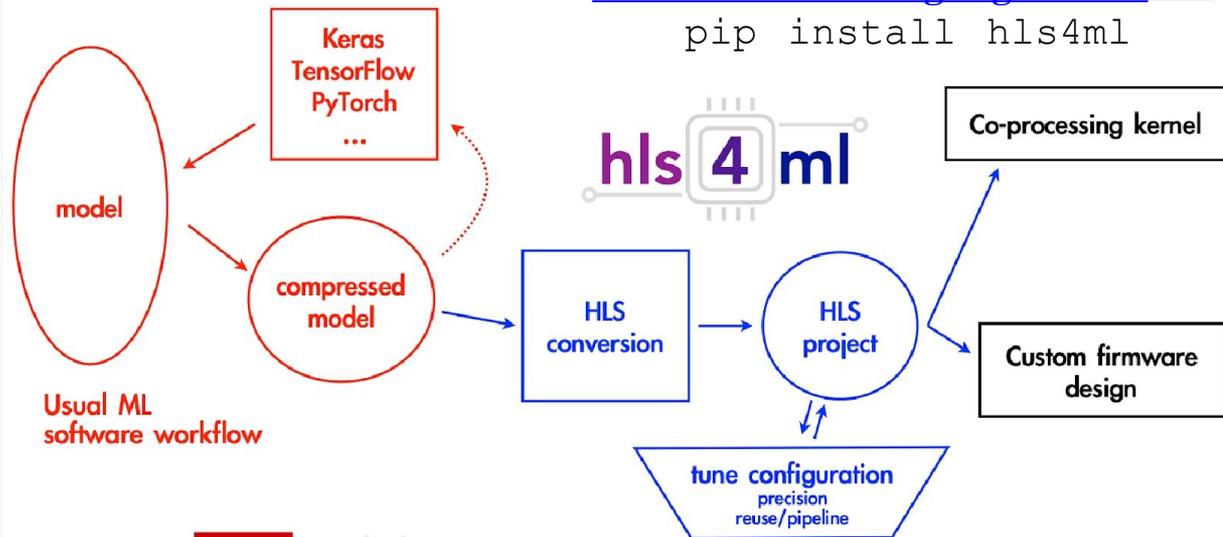


Point cloud: graph

# high level synthesis for machine learning

[fastmachinelearning.org/hls4ml/](https://fastmachinelearning.org/hls4ml/)

```
pip install hls4ml
```





# EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE EuCAIFCon 2024

## Background

Within the Compact Muon Solenoid (CMS) Collaboration, various Deep Neural Networks (DNNs) and Machine Learning (MLs) approaches have been employed to investigate the production of a new massive particle that undergoes decay into Higgs Boson pairs (HH) which further decay into a pair of b-quarks and a pair of tau leptons and discriminate the HH signal from the backgrounds.

However, the mentioned models which are employed are often **complex** and considered black boxes, making it challenging to **interpret** how the task was performed and the data analysis review process.

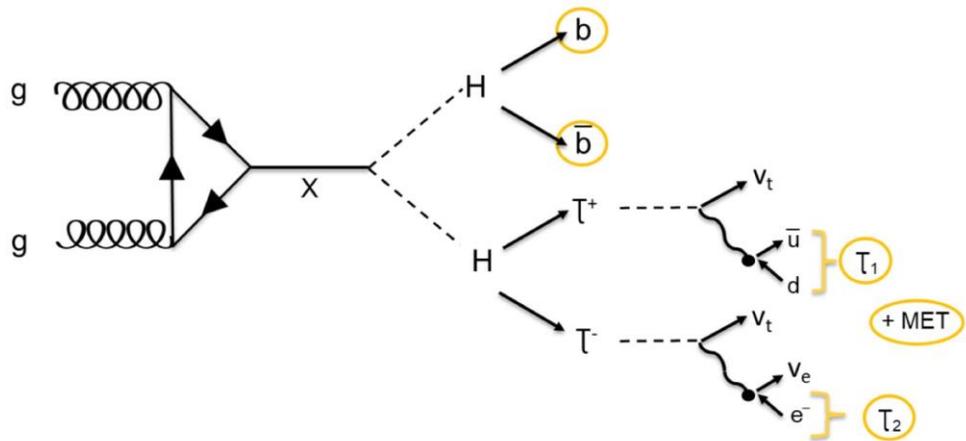


Fig. 1: Decay process outline

## Aim of the work

This work aimed therefore to provide a **better understanding** of how the models work by validating an established **Explainable Artificial Intelligence (XAI)** technique such as SHapley Additive exPlanations (SHAP) [1], aiming for **more interpretable, trustworthy models** and predictions.

$$\begin{aligned}
& -\frac{1}{2}\partial_\nu g_\mu^a \partial_\nu g_\mu^a - g_s f^{abc} \partial_\mu g_\nu^a g_\mu^b g_\nu^c - \frac{1}{4}g_s^2 f^{abc} f^{ade} g_\mu^b g_\nu^c g_\mu^d g_\nu^e + \\
& \frac{1}{2}ig_s^2(\bar{q}_i^\sigma \gamma^\mu q_j^\sigma)g_\mu^a + \bar{G}^a \partial^2 G^a + g_s f^{abc} \partial_\mu \bar{G}^a G^b g_\mu^c - \partial_\nu W_\mu^+ \partial_\nu W_\mu^- - \\
& M^2 W_\mu^+ W_\mu^- - \frac{1}{2}\partial_\nu Z_\mu^0 \partial_\nu Z_\mu^0 - \frac{1}{2c_w^2}M^2 Z_\mu^0 Z_\mu^0 - \frac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu - \frac{1}{2}\partial_\mu H \partial_\mu H - \\
& \frac{1}{2}m_h^2 H^2 - \partial_\mu \phi^+ \partial_\mu \phi^- - M^2 \phi^+ \phi^- - \frac{1}{2}\partial_\mu \phi^0 \partial_\mu \phi^0 - \frac{1}{2c_w^2}M\phi^0 \phi^0 - \beta_h[\frac{2M^2}{g^2} + \\
& \frac{2M}{g}H + \frac{1}{2}(H^2 + \phi^0 \phi^0 + 2\phi^+ \phi^-)] + \frac{2M^4}{g^2}\alpha_h - igc_w[\partial_\nu Z_\mu^0(W_\mu^+ W_\nu^- - \\
& W_\nu^+ W_\mu^-) - Z_\nu^0(W_\mu^+ \partial_\nu W_\mu^- - W_\mu^- \partial_\nu W_\mu^+) + Z_\mu^0(W_\nu^+ \partial_\nu W_\mu^- - \\
& W_\nu^- \partial_\nu W_\mu^+)] - ig_s w[\partial_\nu A_\mu(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) - A_\nu(W_\mu^+ \partial_\nu W_\mu^- - \\
& W_\mu^- \partial_\nu W_\mu^+) + A_\mu(W_\nu^+ \partial_\nu W_\mu^- - W_\nu^- \partial_\nu W_\mu^+)] - \frac{1}{2}g^2 W_\mu^+ W_\mu^- W_\nu^+ W_\nu^- + \\
& \frac{1}{2}g^2 W_\mu^+ W_\nu^- W_\mu^+ W_\nu^- + g^2 c_w^2(Z_\mu^0 W_\mu^+ Z_\nu^0 W_\nu^- - Z_\mu^0 Z_\mu^0 W_\nu^+ W_\nu^-) + \\
& g^2 s_w^2(A_\mu W_\mu^+ A_\nu W_\nu^- - A_\mu A_\mu W_\nu^+ W_\nu^-) + g^2 s_w c_w[A_\mu Z_\nu^0(W_\mu^+ W_\nu^- - \\
& W_\nu^+ W_\mu^-) - 2A_\mu Z_\mu^0 W_\nu^+ W_\nu^-] - g\alpha[H^3 + H\phi^0 \phi^0 + 2H\phi^+ \phi^-] - \\
& \frac{1}{8}g^2 \alpha_h[H^4 + (\phi^0)^4 + 4(\phi^+ \phi^-)^2 + 4(\phi^0)^2 \phi^+ \phi^- + 4H^2 \phi^+ \phi^- + 2(\phi^0)^2 H^2] - \\
& gM W_\mu^+ W_\mu^- H - \frac{1}{2}g\frac{M}{c_w^2}Z_\mu^0 Z_\mu^0 H - \frac{1}{2}ig[W_\mu^+(\phi^0 \partial_\mu \phi^- - \phi^- \partial_\mu \phi^0) - \\
& W_\mu^-(\phi^0 \partial_\mu \phi^+ - \phi^+ \partial_\mu \phi^0)] + \frac{1}{2}g[W_\mu^+(H \partial_\mu \phi^- - \phi^- \partial_\mu H) - W_\mu^-(H \partial_\mu \phi^+ - \\
& \phi^+ \partial_\mu H)] + \frac{1}{2}g\frac{1}{c_w}(Z_\mu^0(H \partial_\mu \phi^0 - \phi^0 \partial_\mu H) - ig\frac{s_w^2}{c_w}MZ_\mu^0(W_\mu^+ \phi^- - W_\mu^- \phi^+) + \\
& ig_s wMA_\mu(W_\mu^+ \phi^- - W_\mu^- \phi^+) - ig\frac{1-2c_w^2}{2c_w}Z_\mu^0(\phi^+ \partial_\mu \phi^- - \phi^- \partial_\mu \phi^+) + \\
& ig_s wA_\mu(\phi^+ \partial_\mu \phi^- - \phi^- \partial_\mu \phi^+) - \frac{1}{4}g^2 W_\mu^+ W_\mu^- [H^2 + (\phi^0)^2 + 2\phi^+ \phi^-] - \\
& \frac{1}{4}g^2 \frac{1}{c_w^2}Z_\mu^0 Z_\mu^0 [H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2 \phi^+ \phi^-] - \frac{1}{2}g^2 \frac{s_w^2}{c_w}Z_\mu^0 \phi^0 (W_\mu^+ \phi^- + \\
& W_\mu^- \phi^+) - \frac{1}{2}ig^2 \frac{s_w^2}{c_w}Z_\mu^0 H (W_\mu^+ \phi^- - W_\mu^- \phi^+) + \frac{1}{2}g^2 s_w A_\mu \phi^0 (W_\mu^+ \phi^- + \\
& W_\mu^- \phi^+) + \frac{1}{2}ig^2 s_w A_\mu H (W_\mu^+ \phi^- - W_\mu^- \phi^+) - g^2 \frac{s_w}{c_w}(2c_w^2 - 1)Z_\mu^0 A_\mu \phi^+ \phi^- - \\
& g^1 s_w^2 A_\mu A_\mu \phi^+ \phi^-] - \bar{e}^\lambda (\gamma \partial + m_e^\lambda) e^\lambda - \bar{\nu}^\lambda \gamma \partial \nu^\lambda - \bar{u}_j^\lambda (\gamma \partial + m_u^\lambda) u_j^\lambda - \\
& \bar{d}_j^\lambda (\gamma \partial + m_d^\lambda) d_j^\lambda + ig_s w A_\mu [-(\bar{e}^\lambda \gamma^\mu e^\lambda) + \frac{2}{3}(\bar{u}_j^\lambda \gamma^\mu u_j^\lambda) - \frac{1}{3}(\bar{d}_j^\lambda \gamma^\mu d_j^\lambda)] + \\
& \frac{ig}{4c_w}Z_\mu^0 [(\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{e}^\lambda \gamma^\mu (4s_w^2 - 1 - \gamma^5) e^\lambda) + (\bar{u}_j^\lambda \gamma^\mu (\frac{4}{3}s_w^2 - \\
& 1 - \gamma^5) u_j^\lambda) + (\bar{d}_j^\lambda \gamma^\mu (1 - \frac{8}{3}s_w^2 - \gamma^5) d_j^\lambda)] + \frac{ig}{2\sqrt{2}}W_\mu^+ [(\bar{\nu}^\lambda \gamma^\mu (1 + \gamma^5) e^\lambda) + \\
& (\bar{u}_j^\lambda \gamma^\mu (1 + \gamma^5) C_{\lambda\kappa} d_j^\kappa)] + \frac{ig}{2\sqrt{2}}W_\mu^- [(\bar{e}^\lambda \gamma^\mu (1 + \gamma^5) \nu^\lambda) + (\bar{d}_j^\kappa C_{\lambda\kappa}^\dagger \gamma^\mu (1 + \\
& \gamma^5) u_j^\lambda)] + \frac{ig}{2\sqrt{2}}\frac{m_\lambda^\lambda}{M}[-\phi^+(\bar{\nu}^\lambda (1 - \gamma^5) e^\lambda) + \phi^-(\bar{e}^\lambda (1 + \gamma^5) \nu^\lambda)] - \\
& \frac{g}{2}\frac{m_\lambda^\lambda}{M}[H(\bar{e}^\lambda e^\lambda) + i\phi^0(\bar{e}^\lambda \gamma^5 e^\lambda)] + \frac{ig}{2M\sqrt{2}}\phi^+[-m_\lambda^\kappa(\bar{u}_j^\lambda C_{\lambda\kappa}(1 - \gamma^5) d_j^\kappa) + \\
& m_u^\lambda(\bar{u}_j^\lambda C_{\lambda\kappa}(1 + \gamma^5) d_j^\kappa)] + \frac{ig}{2M\sqrt{2}}\phi^-[m_\lambda^\lambda(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1 + \gamma^5) u_j^\kappa) - m_u^\kappa(\bar{d}_j^\lambda C_{\lambda\kappa}^\dagger(1 - \\
& \gamma^5) u_j^\kappa) - \frac{g}{2}\frac{m_\lambda^\lambda}{M}H(\bar{u}_j^\lambda u_j^\lambda) - \frac{g}{2}\frac{m_\lambda^\lambda}{M}H(\bar{d}_j^\lambda d_j^\lambda) + \frac{ig}{2}\frac{m_\lambda^\lambda}{M}\phi^0(\bar{u}_j^\lambda \gamma^5 u_j^\lambda) - \\
& \frac{ig}{2}\frac{m_\lambda^\lambda}{M}\phi^0(\bar{d}_j^\lambda \gamma^5 d_j^\lambda)] + \bar{X}^+(\partial^2 - M^2)X^+ + \bar{X}^-(\partial^2 - M^2)X^- + \bar{X}^0(\partial^2 - \\
& \frac{M^2}{c_w^2})X^0 + \bar{Y}\partial^2 Y + igc_w W_\mu^+(\partial_\mu \bar{X}^0 X^- - \partial_\mu \bar{X}^+ X^0) + ig_s w W_\mu^+(\partial_\mu \bar{Y} X^- - \\
& \partial_\mu \bar{X}^+ Y) + igc_w W_\mu^-(\partial_\mu \bar{X}^- X^0 - \partial_\mu \bar{X}^0 X^+) + ig_s w W_\mu^-(\partial_\mu \bar{X}^- Y - \\
\end{aligned}$$

Ну и наконец, весьма полезный доклад в контексте тематики данной конференции:  
2024-04-24 CERIAS – «***Defining Trusted Artificial Intelligence for the National Security Space***»  
<https://www.youtube.com/watch?v=dXZMtBFg-U>

... там есть любопытные акценты на трендах т.н. «доверенности» ИИ в части ИБ нюансов...