



Istituto Nazionale di Fisica Nucleare  
Sezione di Bologna



# Intelligenza Artificiale e INFN: quale futuro?

Daniele Bonacorsi

65esimo anniversario della Sezione INFN-Bologna - 20 Dicembre 2021, Bologna

# Calcolo scientifico nella fisica INFN

Il **calcolo scientifico** è una tecnologia abilitante per la missione scientifica dell'INFN

- non è stato finora un fattore limitante per la fisica

Affinché questo resti vero, nei prossimi decenni abbiamo grosse sfide.

Sfide note: il computing resta un significativo cost driver

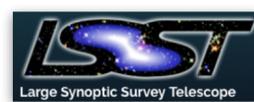
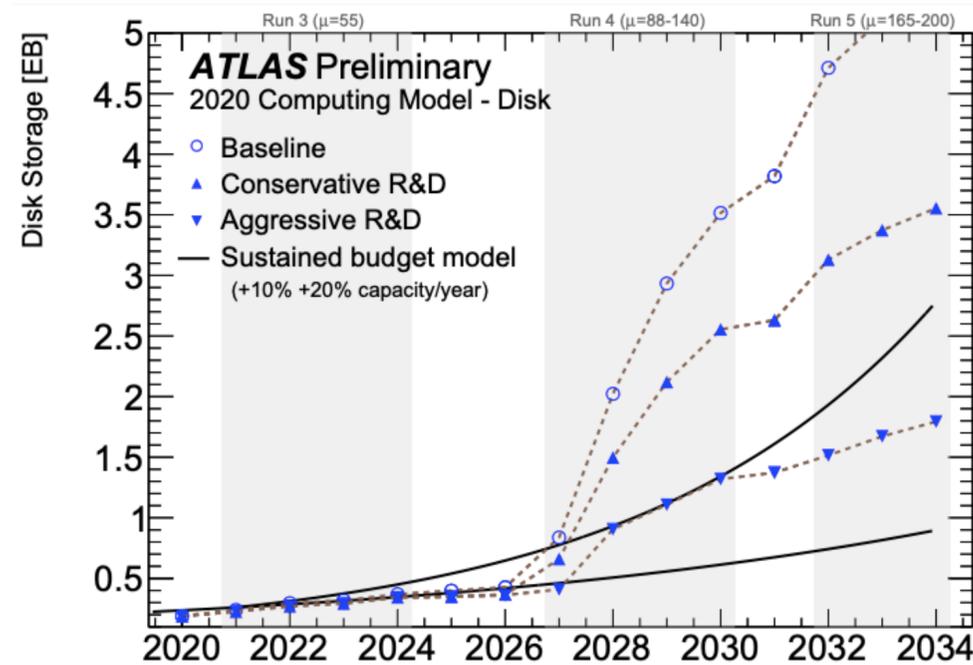
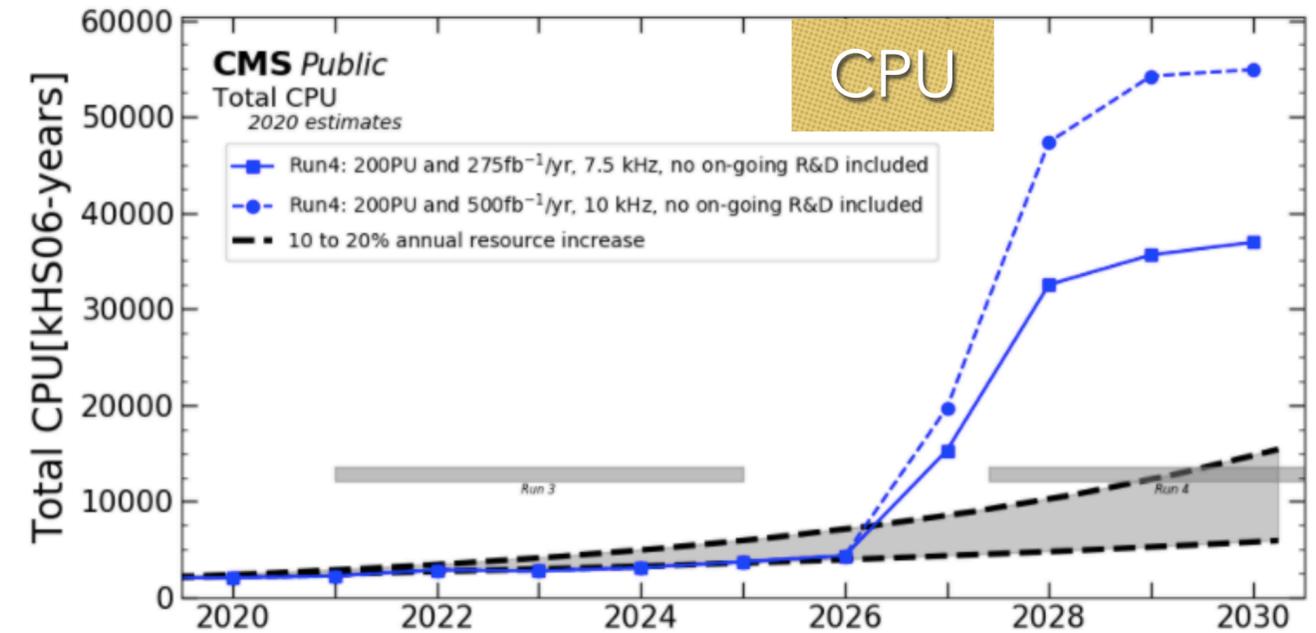
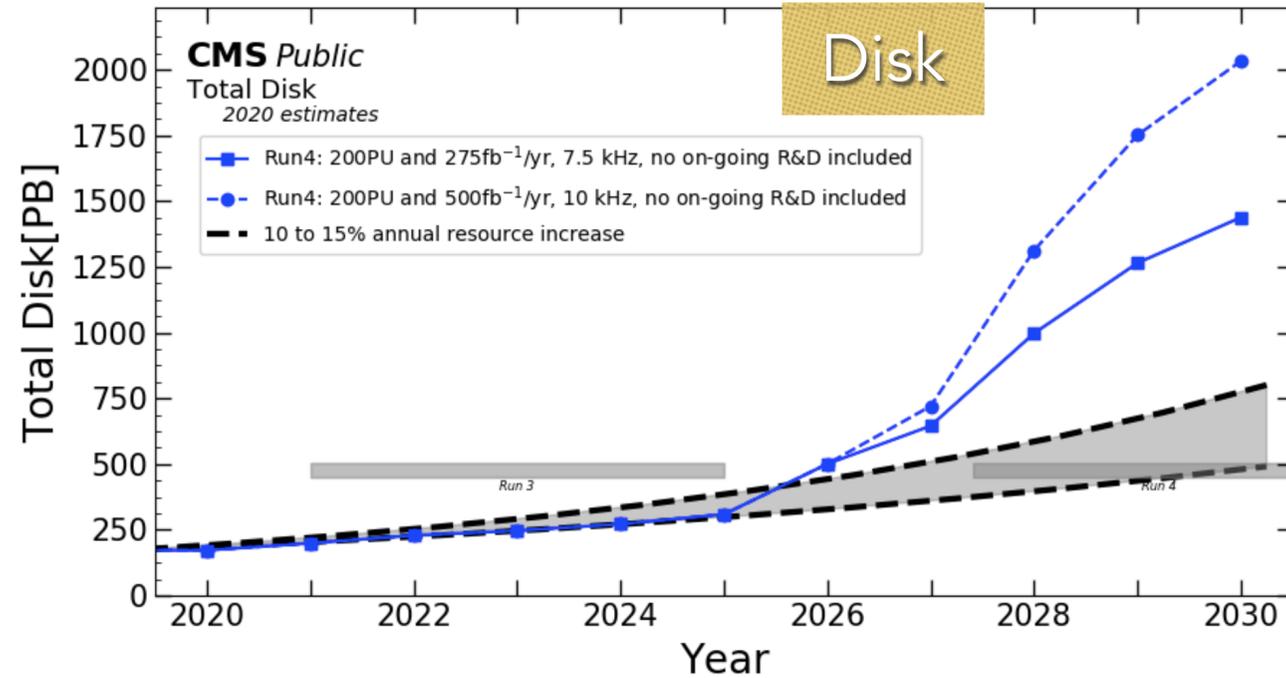
- → evoluzioni adiabatiche dei modelli di calcolo, costanti ottimizzazioni nell'uso delle risorse, etc

Sfide nuove: crescita nella richiesta di risorse e nelle prestazioni necessarie

- e.g. HL-LHC, astro-particle, theory, .. e nuovi esperimenti
- complessità ulteriore, legata all'incertezza nella definizione quantitativa delle necessità, e la specificazione di un ambiente di calcolo adeguato
- → ulteriori ottimizzazioni, ma anche valutazione ed eventuale adozione di paradigmi anche rivoluzionari

# La scala delle sfide in HEP (non solo INFN)

Disclaimer: non una lista completa. e solo su fisica sperimentale



- 10x trigger rate, 6x event complexity (+ detector) >60x resources needs (soprattutto disk)
- ~50-300 PB/yr: major challenge on soft/comp and data movement
- multi-messenger astronomy: processing velocity come challenge

NB: non solo più eventi ma anche rivelatori ad alta-granularità: HEP pagherà il prezzo di non aver sempre considerato i costi del computing fin dalla fase di design

<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ComputingandSoftwarePublicResults>

<https://twiki.cern.ch/twiki/bin/view/CMSPublic/CMSOfflineComputingResults>

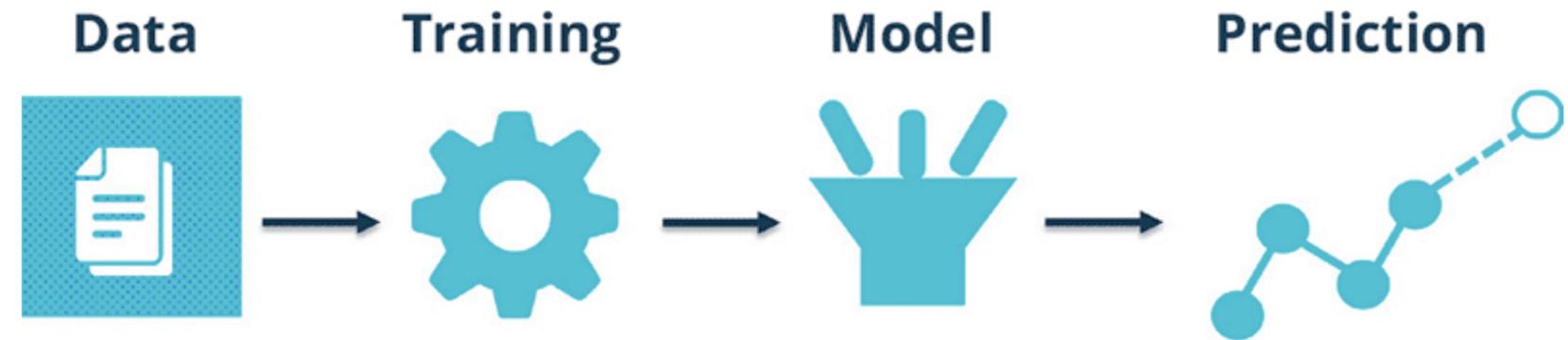
# AI vs ML vs DL

AI = Artificial Intelligence

ML = Machine Learning

NN = Neural Network

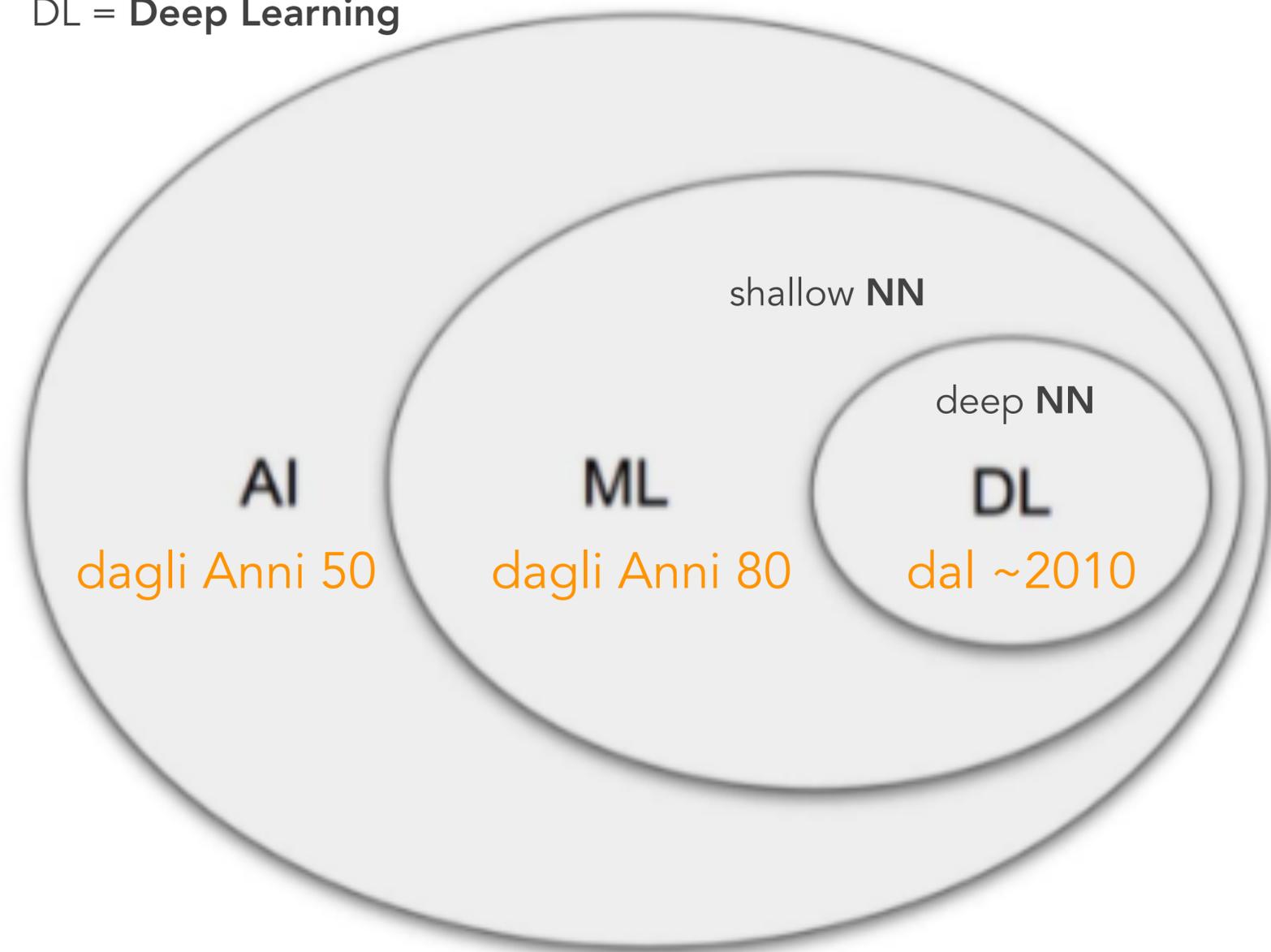
DL = Deep Learning



Goal: estrarre "actionable insight" da (big) data

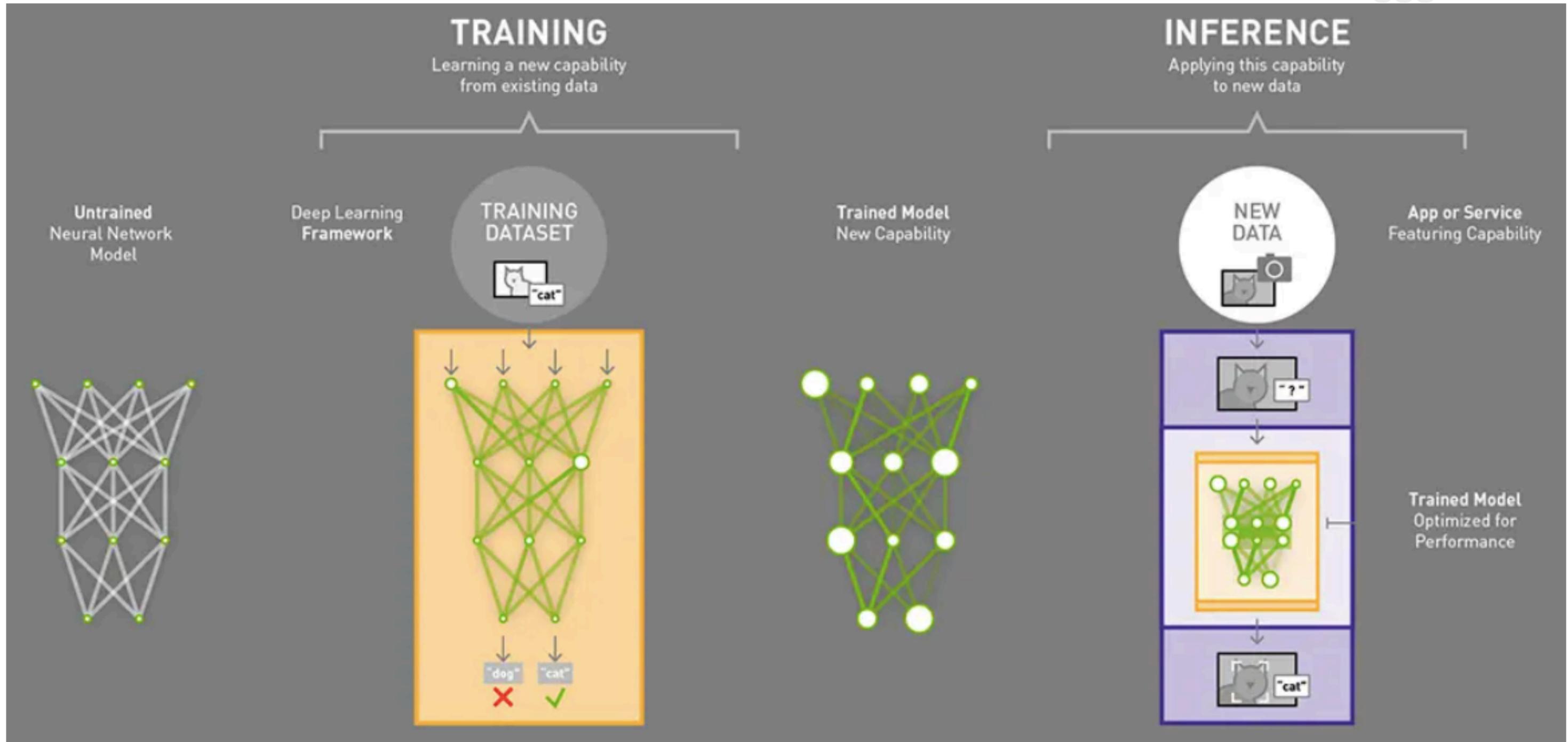
Scelta di un **algoritmo**, e **training** sui dati (attributes vs **features**) per estrarne i **parameters** con tecniche di ottimizzazione (e.g. **gradient descent**) che minimizzano gli errori del modello sulle osservazioni (**cost function**), in un processo governano da tuning di **hyper-parameters**

Risultato: un **ML model** da applicare a nuovi dati

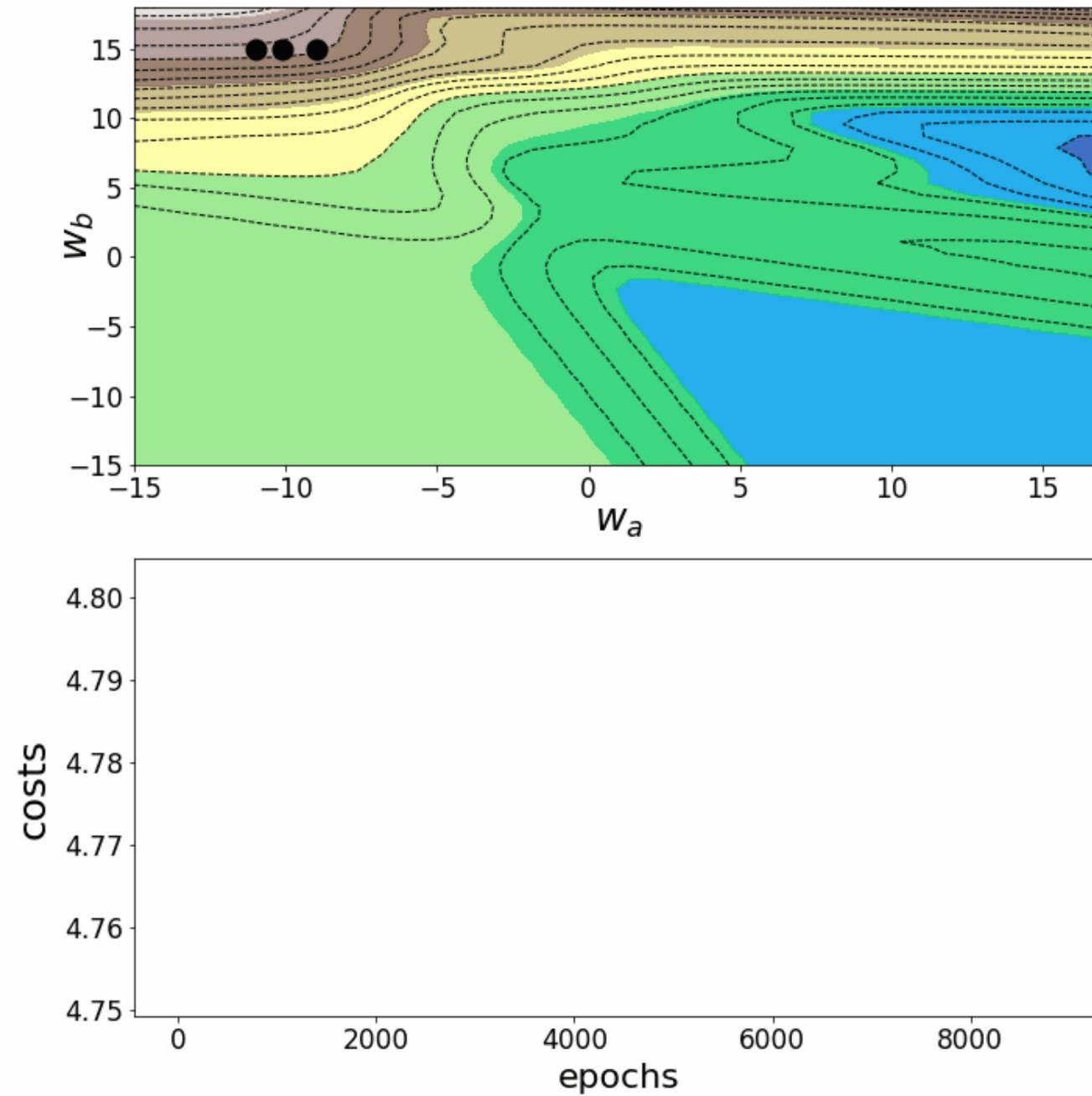
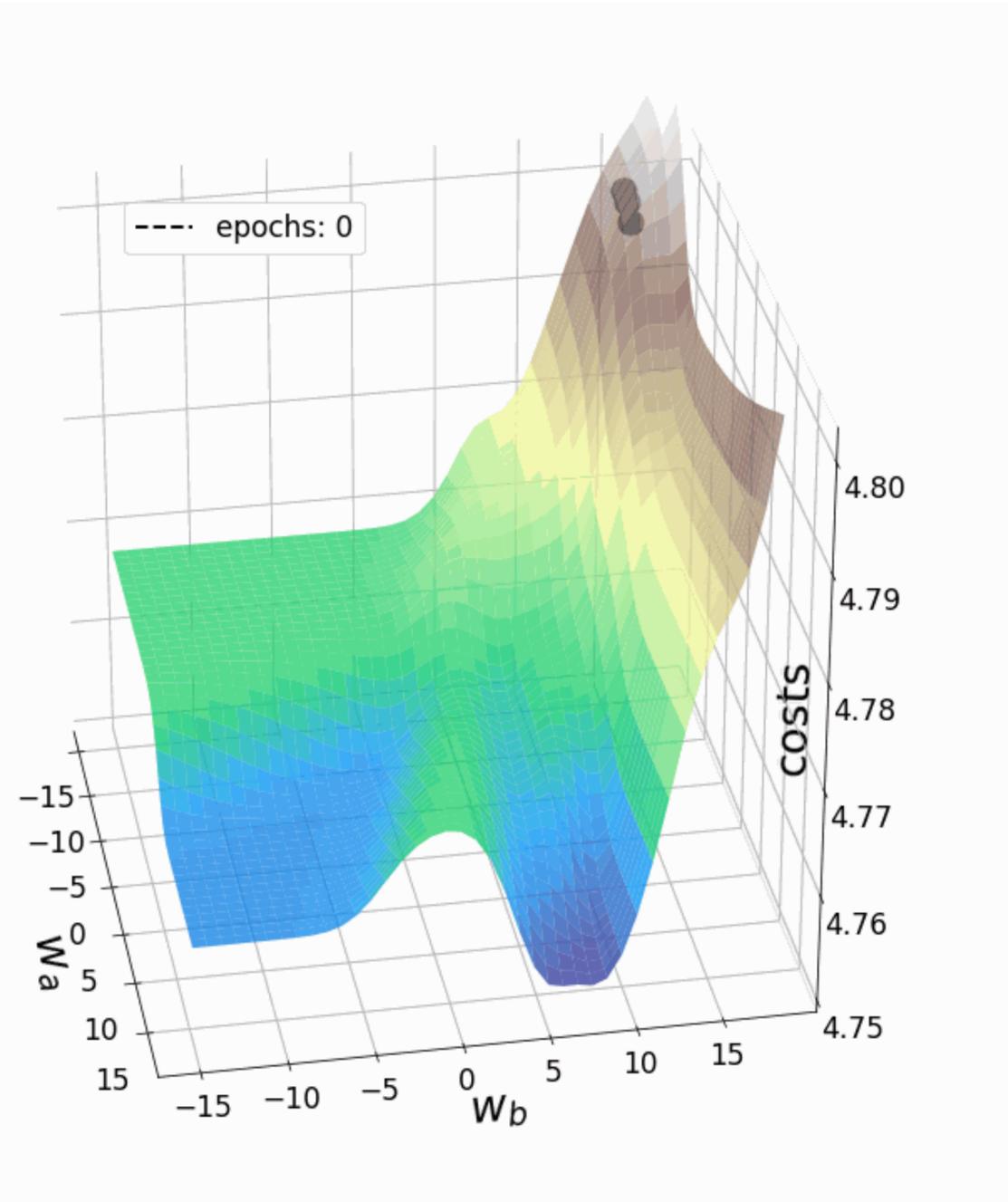


# Allenamento (**training**) e inferenza (**inference**)

SSS

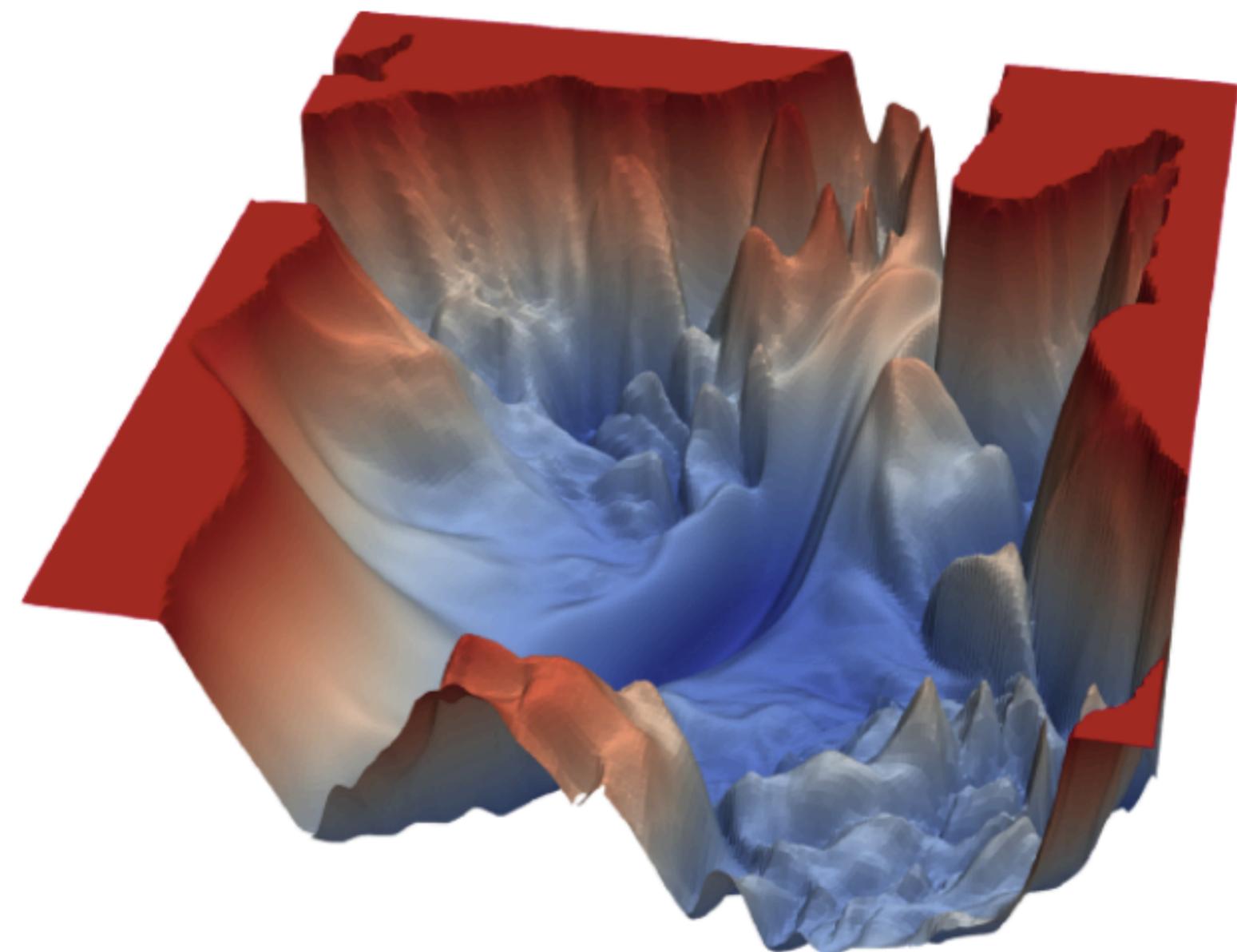


# La ricerca di un minimo di una cost function.. (segue)

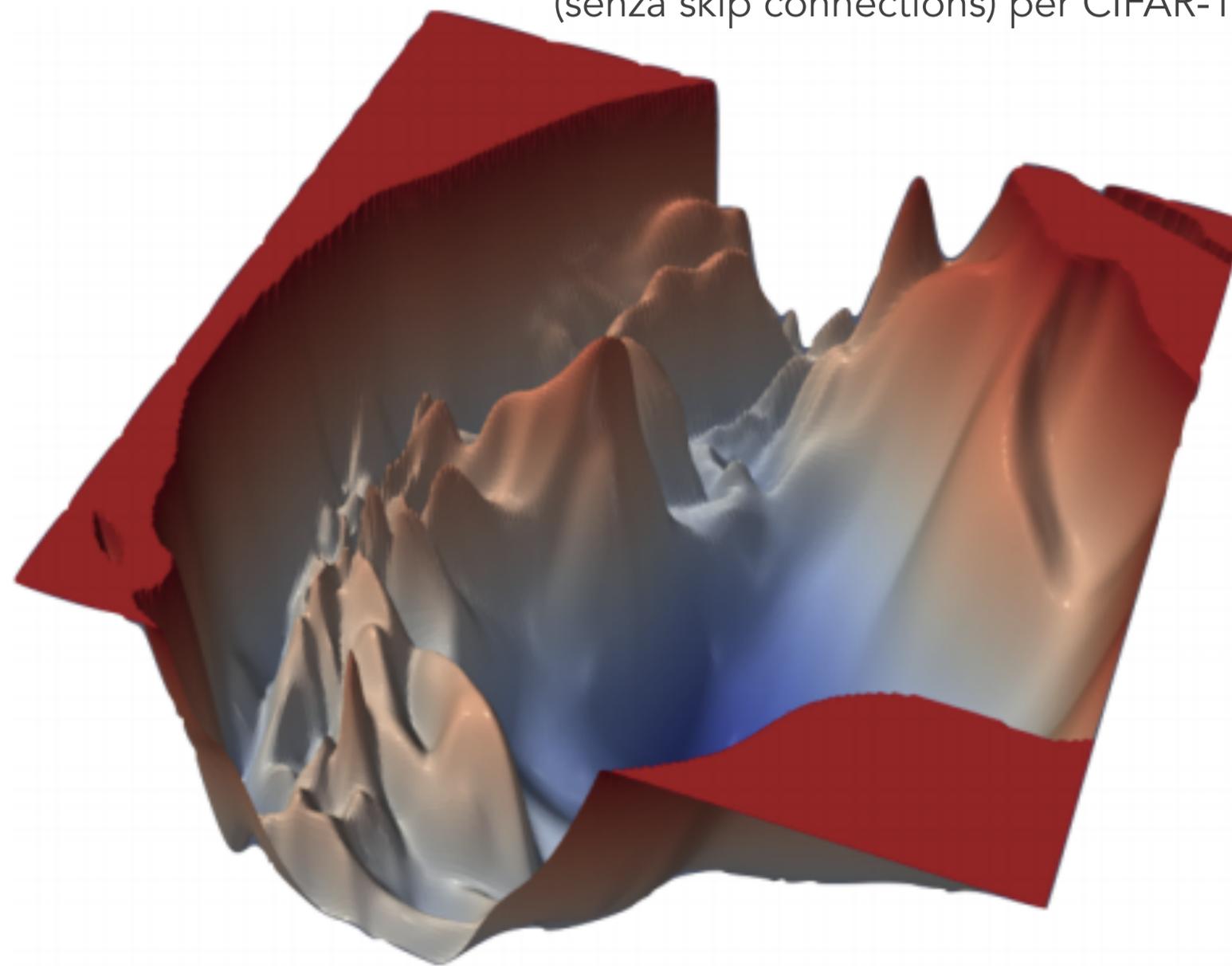


# .. può diventare una sfida in scenari reali di "loss landscape"

Loss landscape di una **VGG-56**  
(una CNN con 56 layers)



Loss landscape di una **ResNet-110**  
(senza skip connections) per CIFAR-10



# Che ruolo deve avere AI nelle sfide dell'INFN?

“La soluzione” vs “una parte della soluzione”

- resistere strenuamente all'hype dell'AI: AI non può essere usata per tutto (e non deve)
- le parti computazionalmente più intense dei workflow HEP sono note → target chiaro

Estrarre un set di **sfide focalizzate e specifiche** dallo spazio dei problemi in HEP e mapparli a idee

- es. modelli discriminativi vs generativi

**Dati** → approcci data-driven modelling sono perseguibili

- “Big” Data (real + simulated) che conosciamo bene → benzina per ML training

**Risorse di calcolo** → high-performance computing infrastructures

- design, deployment, operazione di sistemi di calcolo su larga scala (es. WLCG) → verso ML ops “hyper-fast”, “low-power”, ..

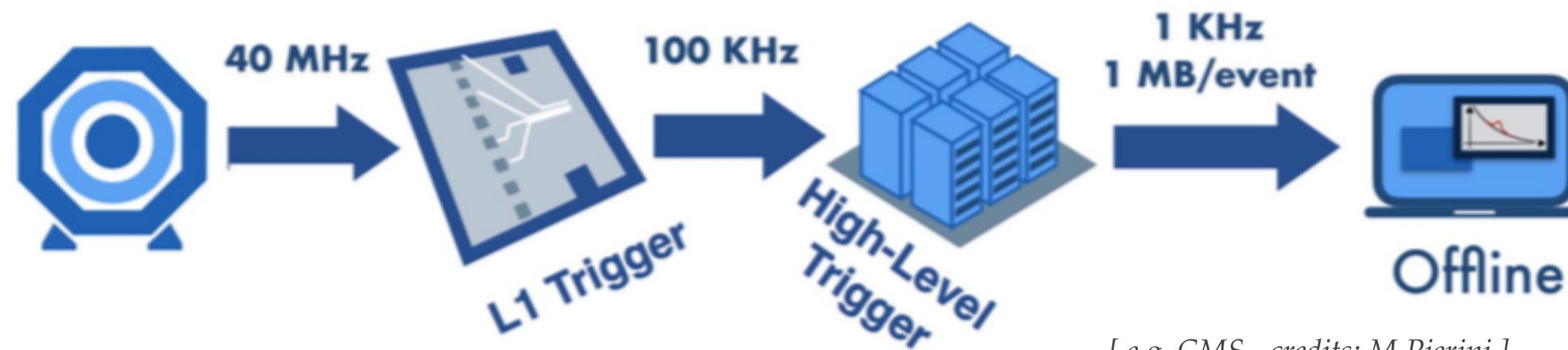
Crescente **sinergia con realtà non-HEP**

- formulazione di problemi HEP in modi comprensibili a ricercatori ML, partecipazione a conferenze ML/DL, ..
- fortissima trazione in HEO software/computing dalle nuove generazioni

# Una panoramica rapida delle aree in cui usiamo ML [1/N]

## ML usato per **data acquisition** e **trigger**

- Bkg and trigger rate reduction
- Signal specific trigger paths
- Anomaly detection in data taking
- Unsupervised new physics mining



[ e.g. CMS - credits: M.Pierini ]

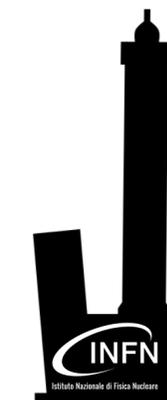
Il trigger di **tutti gli esperimenti HEP** è un forte “driver” per **applicazioni ML ad alte prestazioni**

- sistemi di trigger di prossima generazione → real-time reconstruction → real-time analysis

La sfida è il trade-off tra la **complessità algoritmica** e le **prestazioni** ottenibili sotto **time constraint molto severi** per la fase di inferenza

## Attività di Bologna:

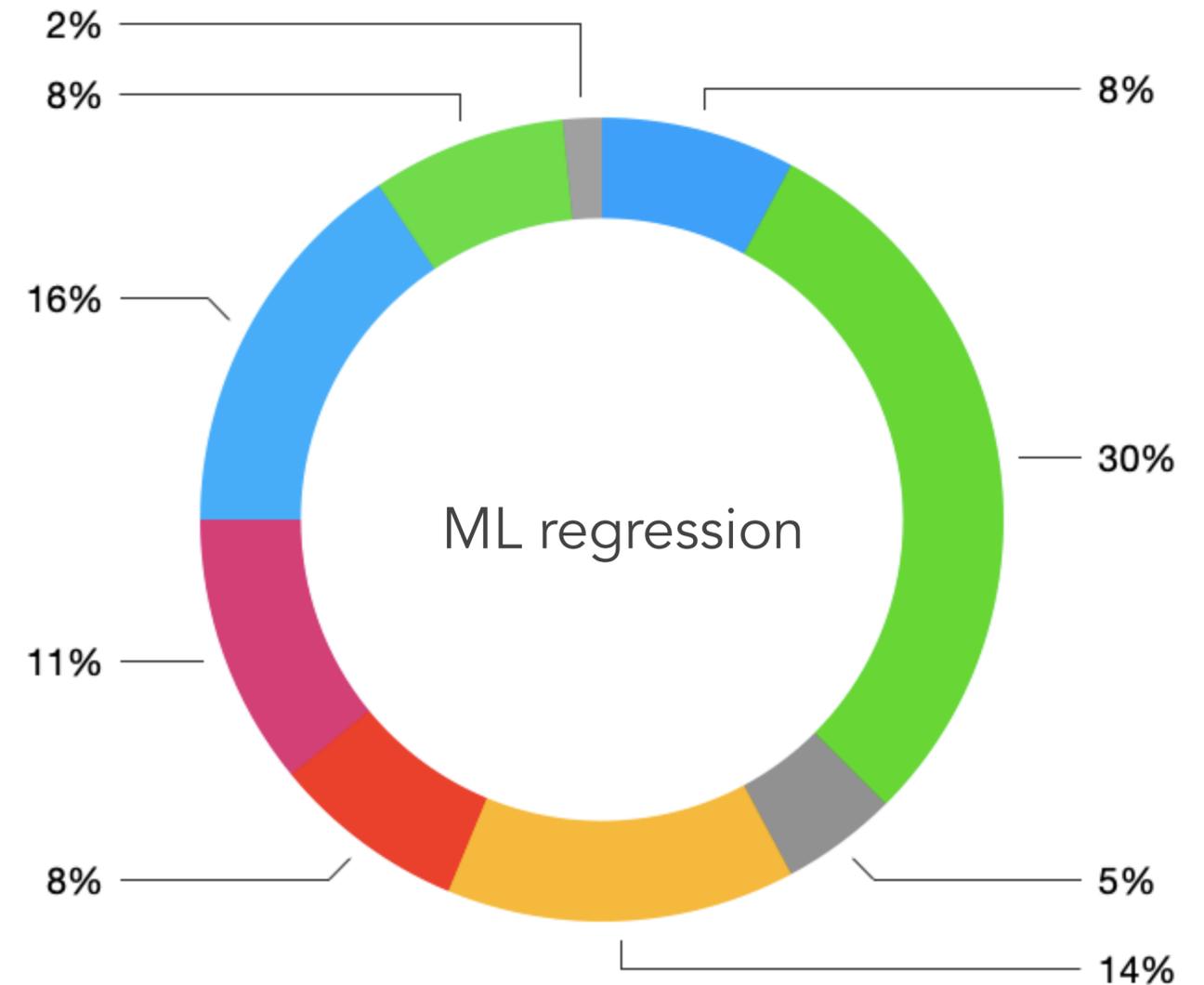
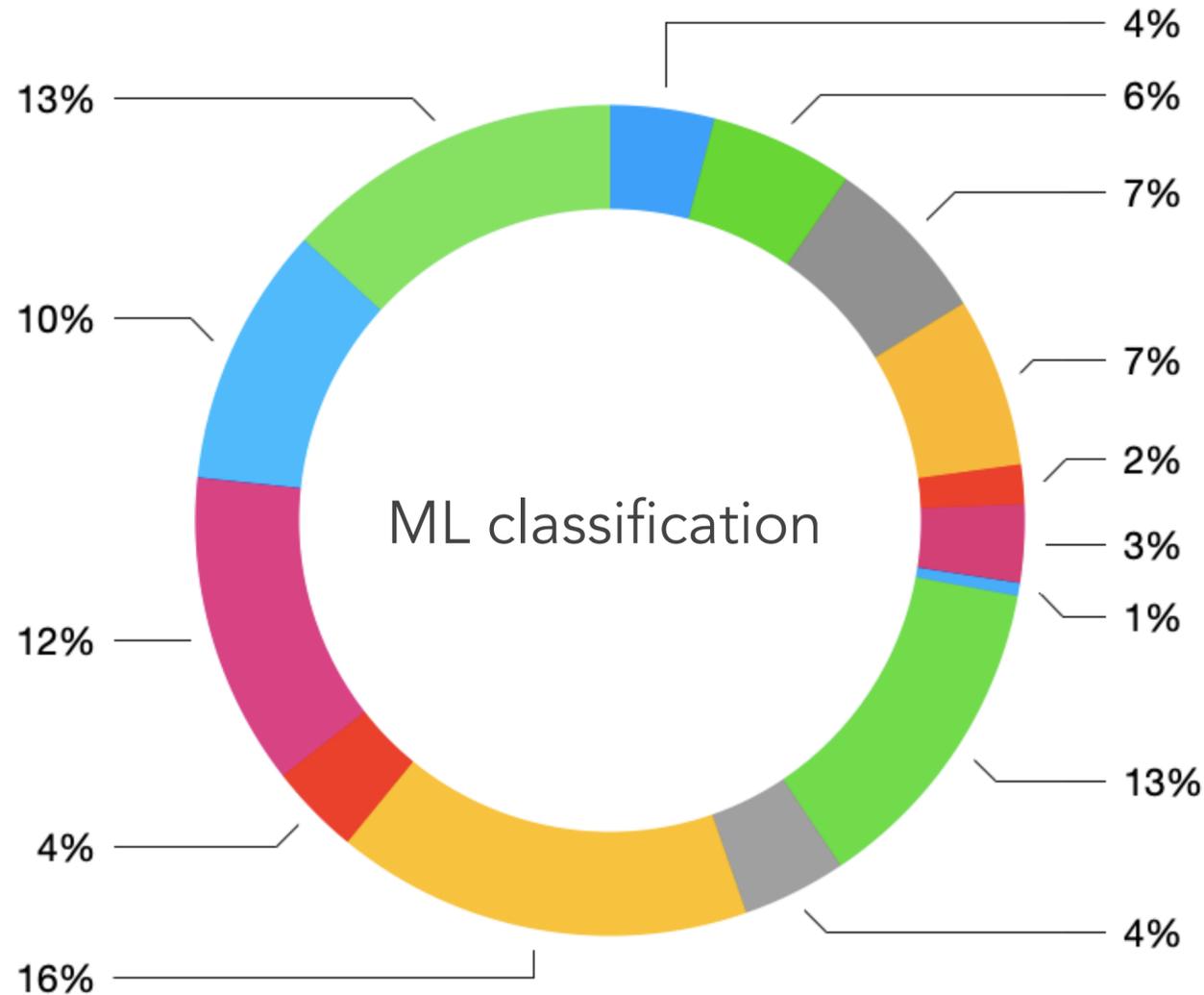
low-latency inference of DL models  
(e.g. CMS Muon Trigger) on FPGA



65<sup>^</sup> INFN-Bologna / D.Bonacorsi

# ML in data analysis

Selezione su dati IML: **171** classification papers, **64** regression papers  
(disclaimer: non realtime, non inclusivo)



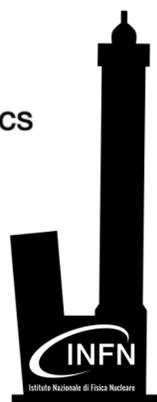
- W/Z tagging
- particle identification
- H→b bar
- neutrino detectors
- quarks and gluons
- direct Dark Matter detectors
- top quark tagging
- cosmology, astroparticle, and cosmic ray physics
- strange jets
- tracking
- b-tagging
- heavy ions / nuclear physics
- flavor physics
- BSM particles and models

**NB: NON** include papers su data representations, learning strategies, technical frameworks per ogni analisi, ..  
(orientativamente altri ~100 papers..)

**Attività di Bologna:**  
reti neurali parametriche, coinvolgimento in progetto ML\_INFN, ..

- Pileup
- Calibration
- Recasting
- Matrix elements
- Parameter estimation
- Parton Distribution Functions (and related)
- Lattice Gauge Theory
- Function Approximation
- Symbolic Regression

65<sup>^</sup> INFN-Bologna / D.Bonacorsi



# Una panoramica rapida delle aree di applicazione di ML [3/N]

## ML in Event **Simulation**

La produzione di eventi simulati (full/fast simulation) è estremamente intensa dal punto di vista delle risorse di calcolo (fino al punto di rischiare di impattare la physics reach degli esperimenti). **ML può aiutare a ridurre tale carico**

- Calorimeter shower surrogate simulator
- Analysis level simulator
- Pile-up overlay generator
- Monte Carlo integration
- ML-enabled fast-simulation
- Invertible full-simulation (probabilistic programming, ...)
- ...

## ML in Event **Reconstruction**

La online/offline reconstruction potrebbe essere in parte **rimpiazzata da modelli surrogati** (approximate → faster) o da **nuovi algoritmi** (che potrebbero dare prestazioni senza precedenti)

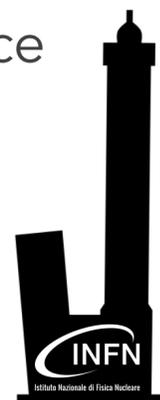
- Charged particle tracking (GraphNN, vertexing, ...)
- Calorimeter reconstruction (local, clustering, ...)
- Particle flow (GraphNN, ...)
- Particle identification (boosted jets, isolation, ...)
- Pileup mitigation
- Energy regression (end-2-end, ...)
- ...

# Una panoramica rapida delle aree di applicazione di ML [4/N]

## ML per le **Operations**

L'applicazione di ML a **non-collision (meta-)data** può aiutare ad **aumentare l'efficienza** e **ridurre la necessità di personpower** nelle Ops, e.g. automatizzando task particolari, sviluppando sistemi intelligenti/adattivi, in ultima analisi agendo per rendere più snella e veloce l'intera catena dalla raccolta dei dati fino all'analisi finale

- Detector control
- Data quality monitoring
- Operational intelligence
- Predictive maintenance
- ...



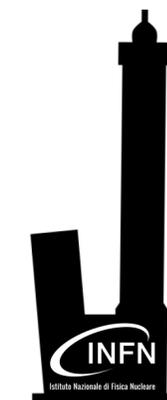
### Attività di Bologna:

Es. real-time anomaly detection in data centers for log-based predictive maintenance using evolving fuzzy-rule-based approach, granular computing, etc

## Methodology in ML

Molto complesso - ma le scienze dure possono fare molto per contribuire direttamente allo sviluppo di una teoria solida alla base dei successivi "empirici" del ML

- Continual learning
- Interpretability and explainability
- Uncertainty quantification
- Incorporating domain knowledge
- Cast into optimisation problems
- ...



### Attività di Bologna:

Collaborazione INFN con Centro Interdipartimentale Alma-AI e con Dottorato in Data Science and Computation di Univ. Bologna

65<sup>^</sup> INFN-Bologna / D.Bonacorsi



# AI, un passato pieno di ostacoli → quale futuro?

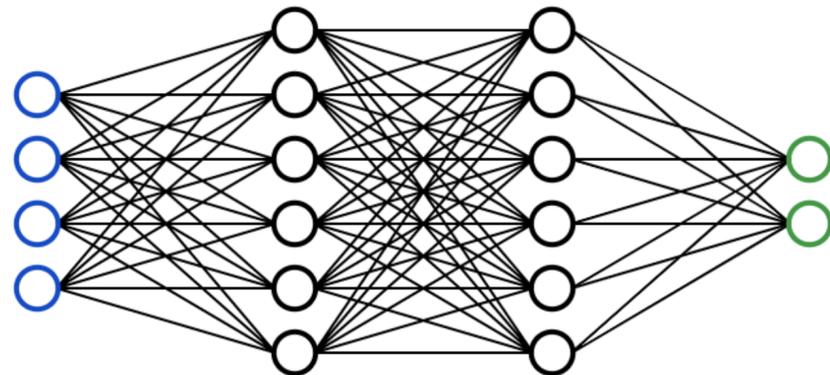
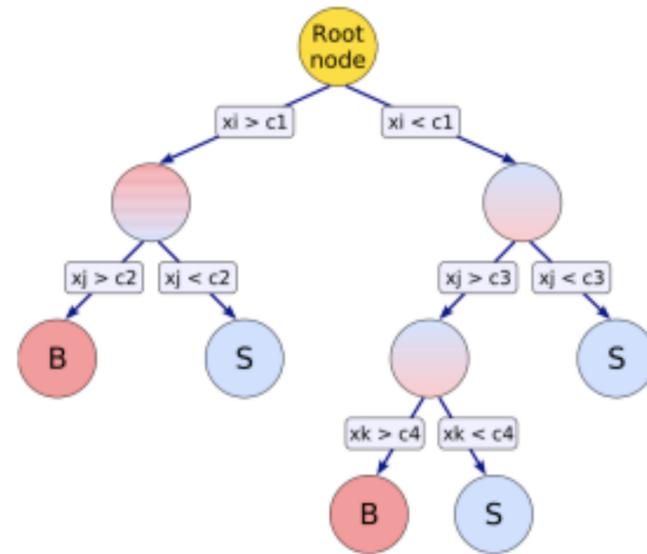
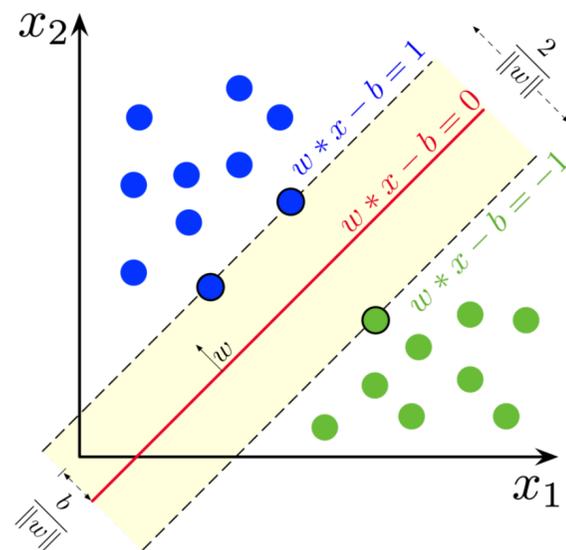
Vari "inverni" dell'AI coniugata come NN:

- Anni '60: shallow neural network (NN) difficili da allenare
- Anni '90: Support Vector Machines (**SVM**), Boosted Decision Trees (**BDT**), ..
- 2000+: architetture avanzate di **NN**



Attuale "primavera" attribuibile a:

- data goes Big
- sviluppo tecnologico e ML research
- accesso Cloud



65<sup>^</sup> INFN-Bologna / D.Bonacorsi

# Community attiva nel definire il ruolo di ML nel futuro di HEP

Computing and Software for Big Science (2019) 3:7  
<https://doi.org/10.1007/s41781-018-0018-8>

ORIGINAL ARTICLE



## A Roadmap for HEP Software and Computing R&D for the 2020s

The HEP Software Foundation<sup>5</sup> · Johannes Albrecht<sup>69</sup> · Antonio Augusto Alves Jr<sup>81</sup> · Guilherme Amadio<sup>5</sup> · Giuseppe Andronico<sup>27</sup> · Nguyen Anh-Ky<sup>122</sup> · Laurent Aphecetche<sup>66</sup> · John Apostolakis<sup>5</sup> · Makoto Asai<sup>63</sup> · Luca Atzori<sup>5</sup> · Marian Babik<sup>5</sup> · Giuseppe Bagliesi<sup>32</sup> · Marilena Bandieramonte<sup>5</sup> · Sunanda Banerjee<sup>16</sup> · Martin Barisits<sup>5</sup> · Lothar A. T. Bauerdick<sup>16</sup> · Stefano Belforte<sup>35</sup> · Douglas Benjamin<sup>82</sup> · Catrin Bernius<sup>63</sup> · Wahid Bhimji<sup>46</sup> · Riccardo Maria Bianchi<sup>105</sup> · Ian Bird<sup>5</sup> · Catherine Biscarat<sup>52</sup> · Jakob Blomer<sup>5</sup> · Kenneth Bloom<sup>97</sup> · Tommaso Boccali<sup>32</sup> · Brian Bockelman<sup>97</sup> · Tomasz Bold<sup>43</sup> · Daniele Bonacorsi<sup>25</sup> · Antonio Boveia<sup>101</sup> · Concezio Bozzi<sup>28</sup> · Marko Bracko<sup>41,93</sup> · David Britton<sup>86</sup> · Andy Buckley<sup>86</sup> · Predrag Buncic<sup>5</sup> · Paolo Calafiura<sup>46</sup>

(...)

### Abstract

Particle physics has an ambitious and broad experimental programme for the coming decades. This programme requires large investments in detector hardware, either to build new facilities and experiments, or to upgrade existing ones. Similarly, it requires commensurate investment in the R&D of software to acquire, manage, process, and analyse the shear amounts of data to be recorded. In planning for the HL-LHC in particular, it is critical that all of the collaborating stakeholders agree on the software goals and priorities, and that the efforts complement each other. In this spirit, this white paper describes the R&D activities required to prepare for this software upgrade.

**Keywords** Particle physics · HL-LHC · Computing & software upgrade · Software performance · Machine learning

## REVIEW

nature

<https://doi.org/10.1038/s41586-018-0361-2>

## Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic<sup>1\*</sup>, Mike Williams<sup>2\*</sup>, David Rousseau<sup>3</sup>, Michael Kagan<sup>4</sup>, Daniele Bonacorsi<sup>5,6</sup>, Alexander Himmel<sup>7</sup>, Adam Aurisano<sup>8</sup>, Kazuhiro Terao<sup>4</sup> & Taritree Wongjirad<sup>9</sup>

Our knowledge of the fundamental particles of nature and their interactions is summarized by the standard model of particle physics. Advancing our understanding in this field has required experiments that operate at ever higher energies and intensities, which produce extremely large and information-rich data samples. The use of machine-learning techniques is revolutionizing how we interpret these data samples, greatly increasing the discovery potential of present and future experiments. Here we summarize the challenges and opportunities that come with the use of machine learning at the frontiers of particle physics.

**T**he standard model of particle physics is supported by an abundance of experimental evidence, yet we know that it cannot be a complete theory of nature because, for example, it cannot incorporate gravity or explain dark matter. Furthermore, many properties of known particles, including neutrinos and the Higgs boson, have not yet been determined experimentally, and the way in which the emergent properties of complex systems of fundamental particles arise from the

### Big data at the LHC

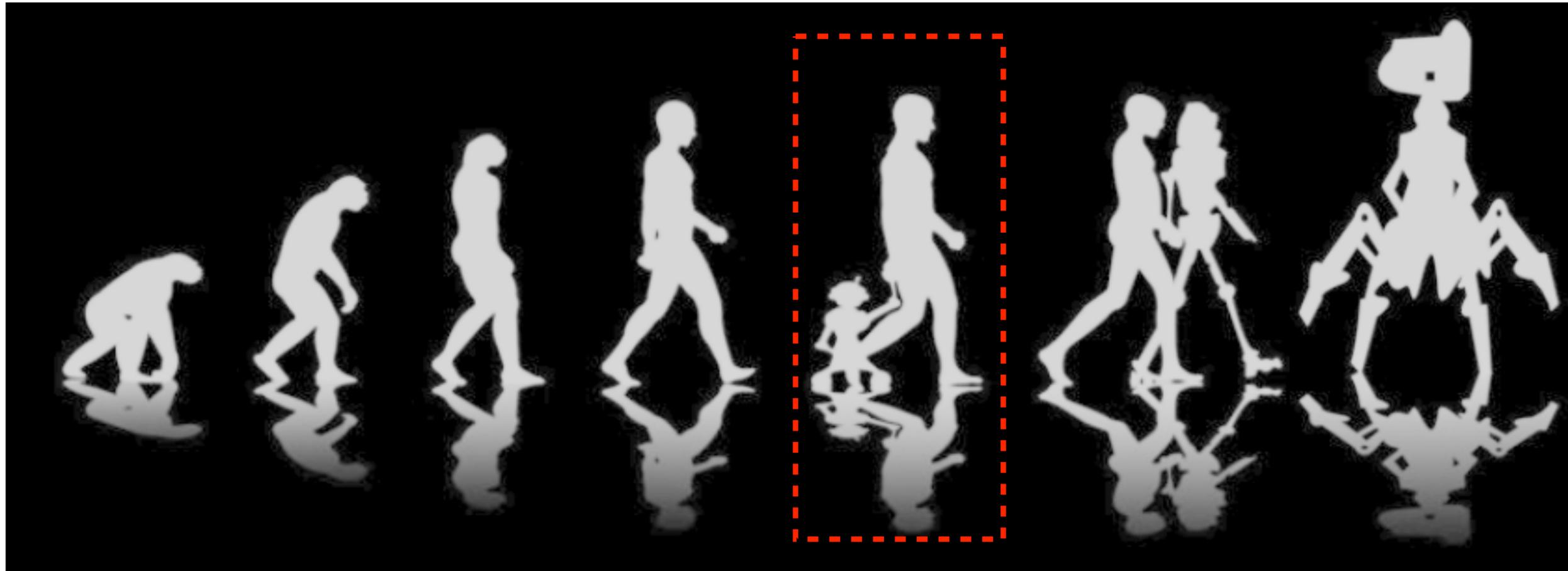
The sensor arrays of the LHC experiments produce data at a rate of about one petabyte per second. Even after drastic data reduction by the custom-built electronics used to readout the sensor arrays, which involves zero suppression of the sparse data streams and the use of various custom compression algorithms, the data rates are still too large to store the data indefinitely—as much as 50 terabytes per second,

[Comp. Softw. Big Sci \(2019\) 3:7](https://doi.org/10.1007/s41781-018-0018-8)

[nature.com/articles/s41586-018-0361-2](https://doi.org/10.1038/s41586-018-0361-2)

Autori anche di  
INFN-Bologna

# Primi passi di ML in HEP: ML "tradizionale"



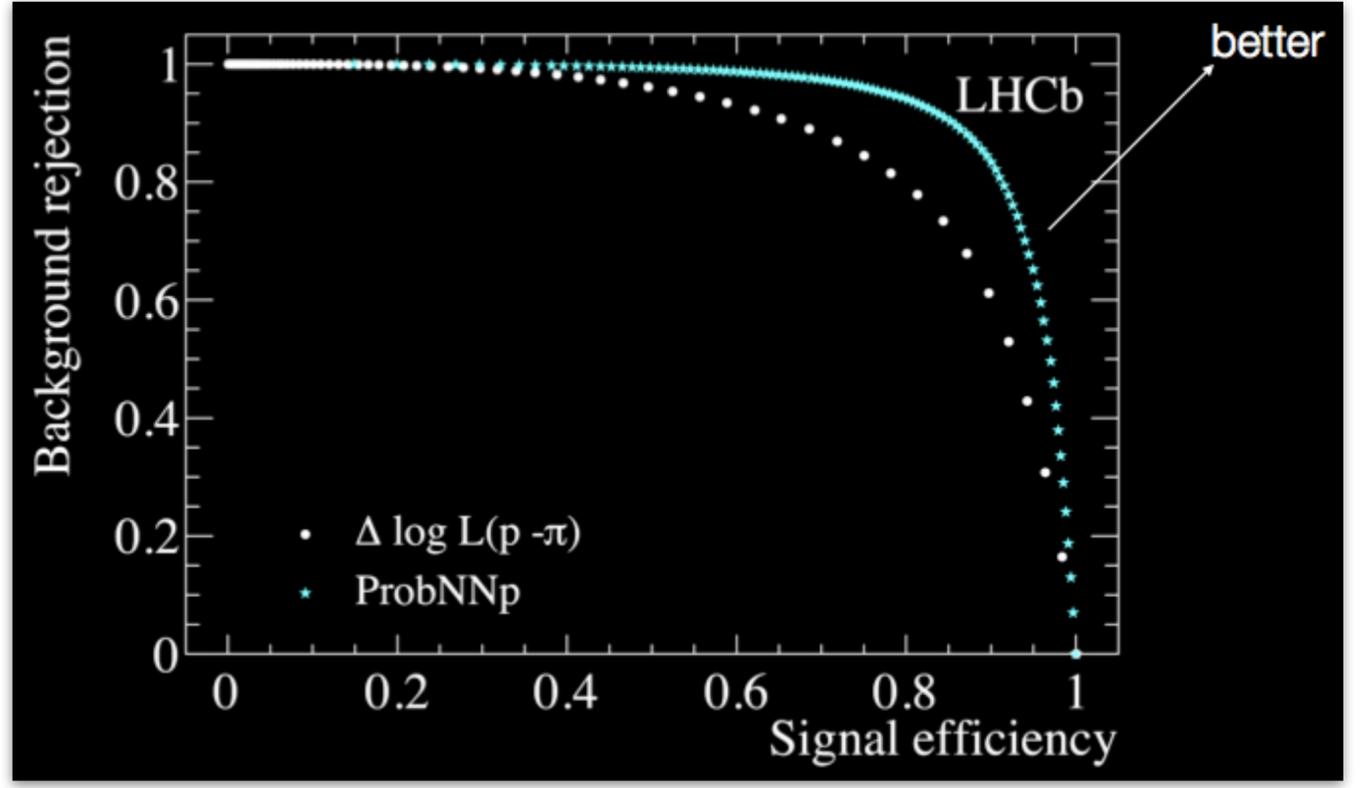
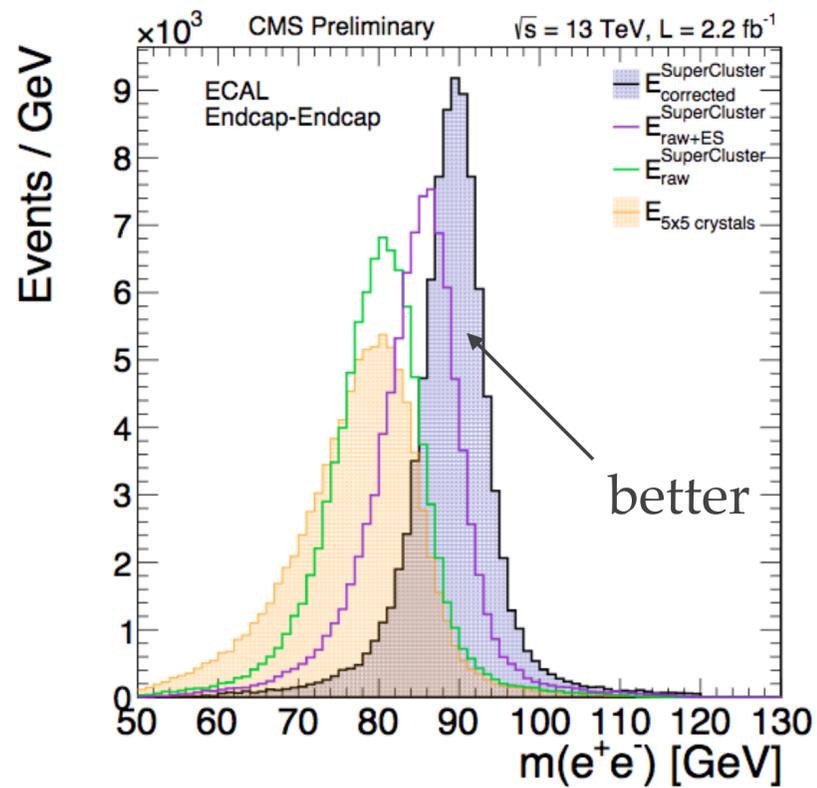
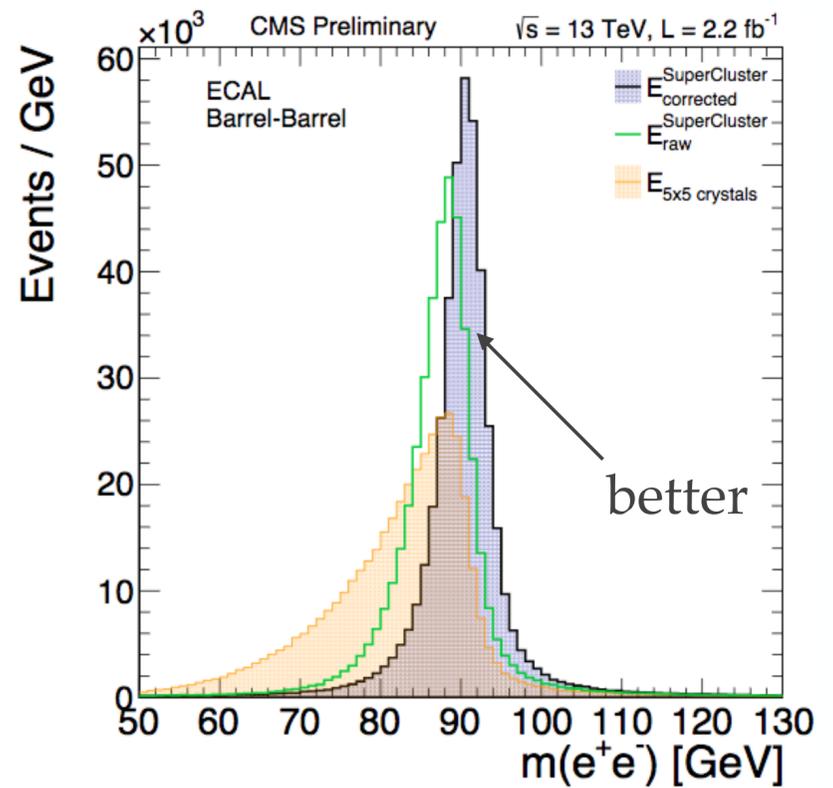
Fino a pochi anni fa, l'uso di ML in HEP si basava sullo sfruttamento di domain knowledge per feature extraction/engineering

- i fisici HEP disegnano high-level features, e le mandano in input a tradizionali algoritmi ML "shallow"

# Particle id, energy resolution, e oltre..

**BDT** usati per il learning delle energy corrections usando tutte le info disponibili nei vari sensori calorimetri

- es. **CMS**: energy sum, recupero bremsstrahlung con supercluster, inclusione della pre-shower energy, poi **energy correction con algo multivariato**



**LHCb** usa delle **NN** allenate su  $O(30)$  features da tutti i sottosistemi, ciascuna allenata a identificare uno specifico tipo di particella

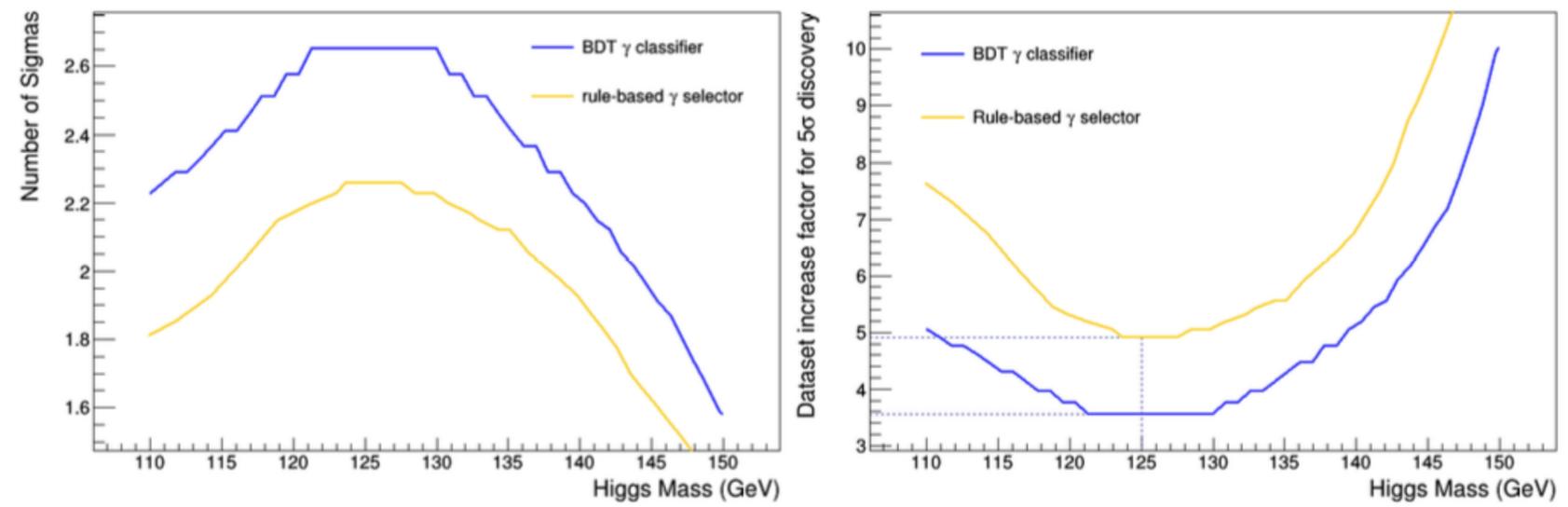
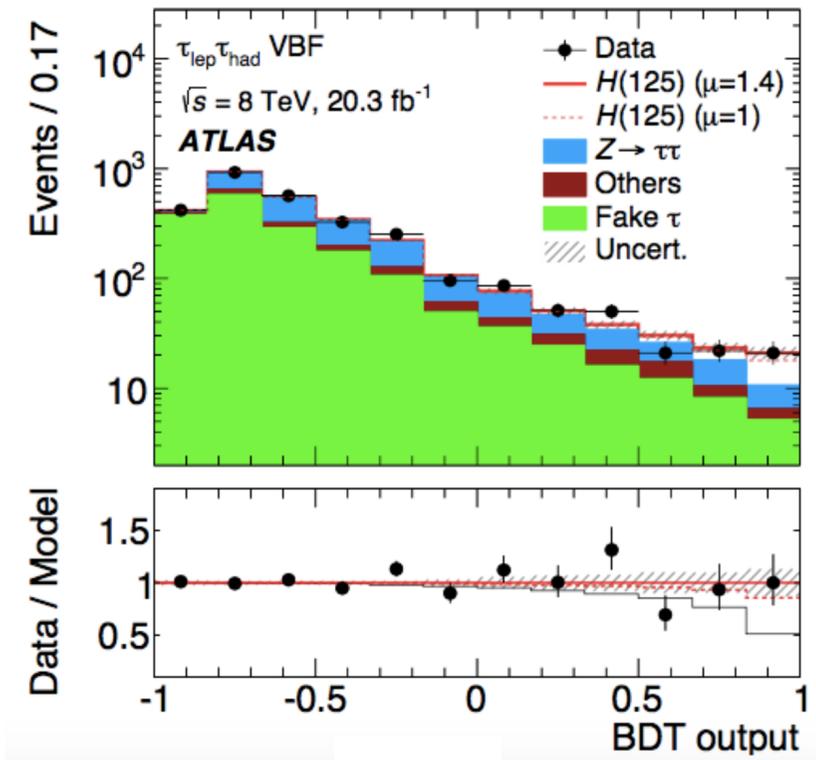
- **mis-ID bkg/particle ridotta di ~3x (e oltre..)**

# Uso di ML in scoperta e studio delle proprietà Higgs

[1] JHEP 04 (2015) 117

Ruolo chiave del ML nella scoperta del bosone di Higgs prima delle attese

- es. diphoton analysis di **CMS**, **BDT** usati per migliorare la risoluzione e selezionare/categorizzare eventi
- → **sensibilità aumentata di un ammontare equivalente al ~50% di dati raccolti in più**



[courtesy M.Pierini]

Studio delle proprietà H: es.  $\tau$  leptons, **ATLAS** su 6 regioni cinematiche distinte, training di una BDT in ciascuna, con 12 features

→ **sensibilità migliorata del ~40% rispetto a un approccio non-ML**

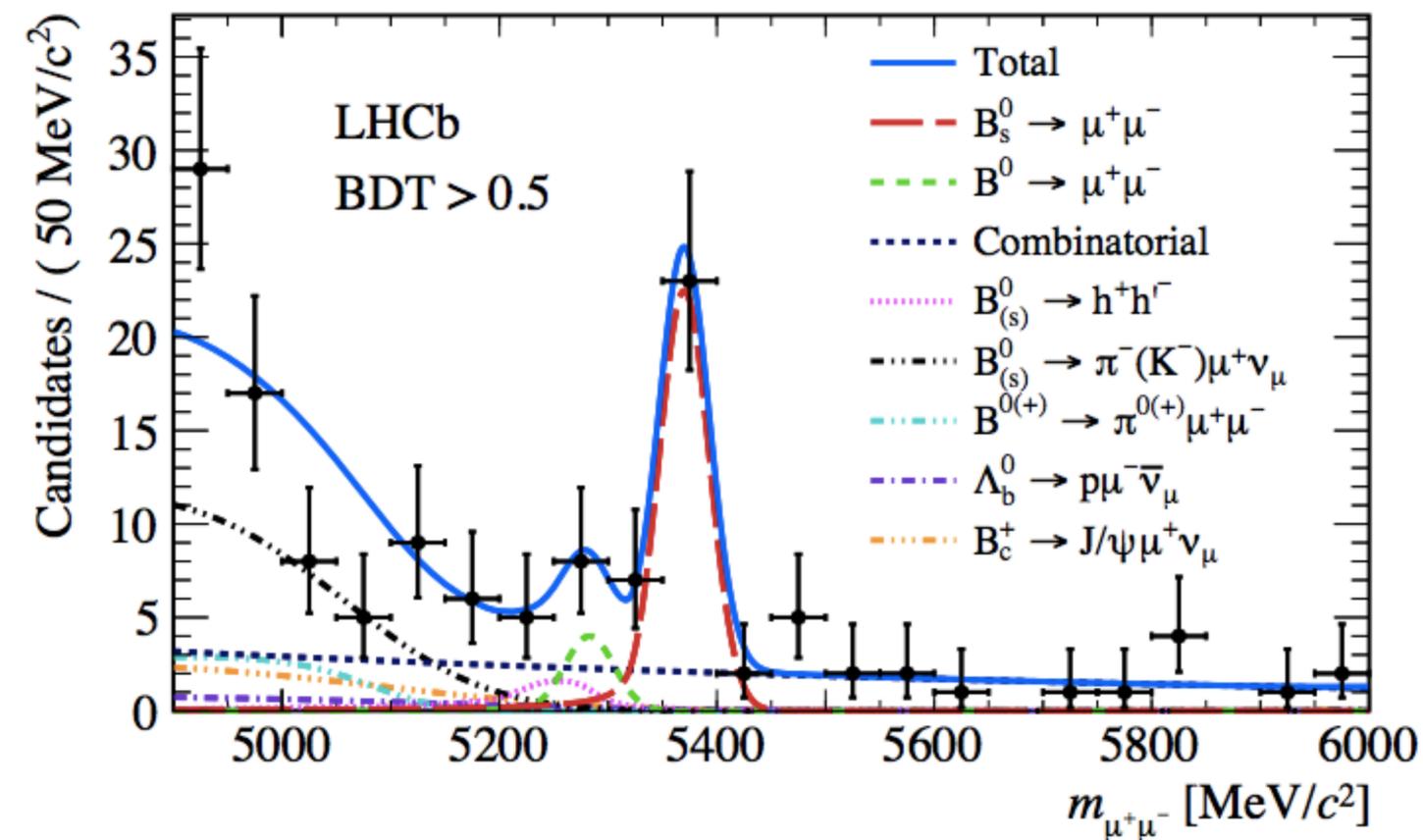
# Test di alta precisione del MS

[1] Nature 522 68–72 (2015)

[2] Phys.Rev.Lett. 118 (2017) 19, 191801

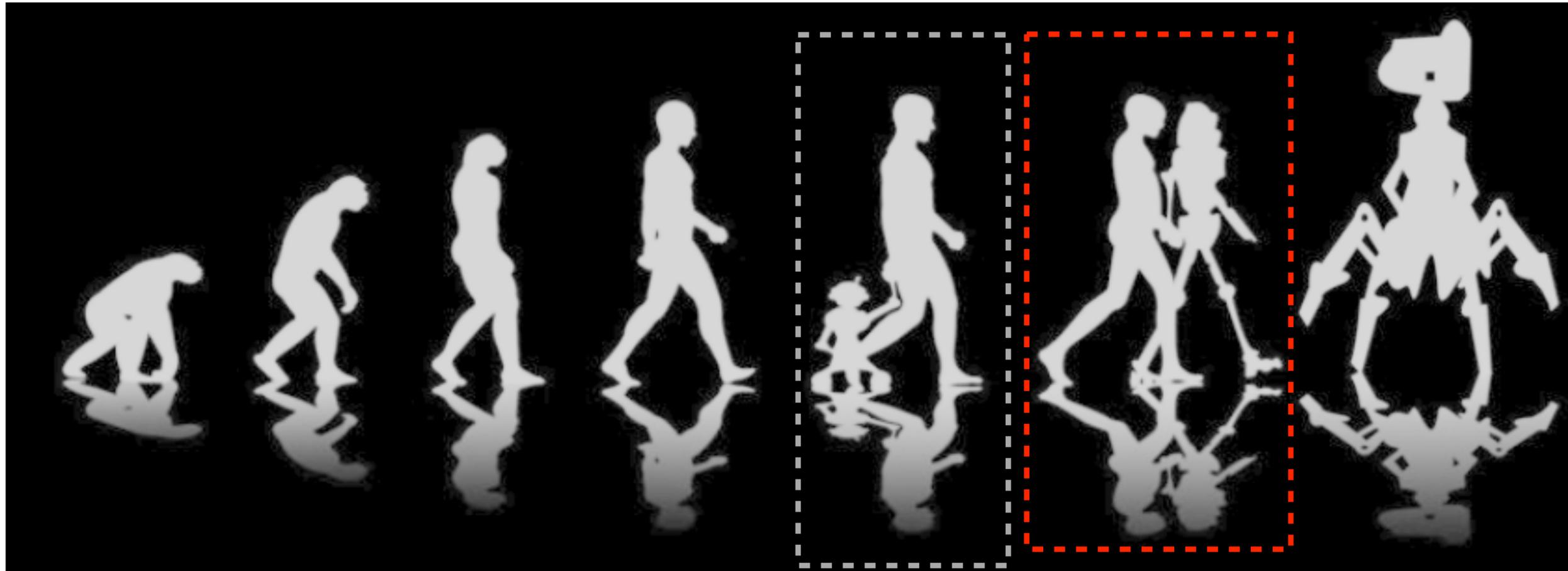
**CMS** e **LHCb**: evidenza per il decadimento raro  $B^0_s \rightarrow \mu^+ \mu^-$  con analisi combinata [1]

- es. **BDT** usati per ridurre la dimensionality dello spazio delle features, poi analisi spettro massa in BDT response bins  $\rightarrow$  decay rate consistente con predizioni SM con precisione  $\sim 25\%$   $\rightarrow$  constraints a estensioni SM
- $\rightarrow$  es. **per avere la stessa sensitività senza ML, LHCb come singolo esperimento avrebbe dovuto raccogliere  $\sim 4x$  più dati**



Mass distribution of the selected  $B^0 \rightarrow \mu^+ \mu^-$  candidates with  $BDT > 0.5$  [2]

# Passi successivi in HEP: oltre il ML tradizionale



Da vari anni, ML in HEP sfrutta tecniche più avanzate, es. NN con architetture evolute di NN

- uso tutto lo spazio delle features a piena dimensionalità per allenare NN profonde (**Deep Learning**), senza feature engineering manuale

→ estrarre il meglio dai dati, a seconda delle tipologie con architetture nuove e più esotiche (come in CV e NLP)

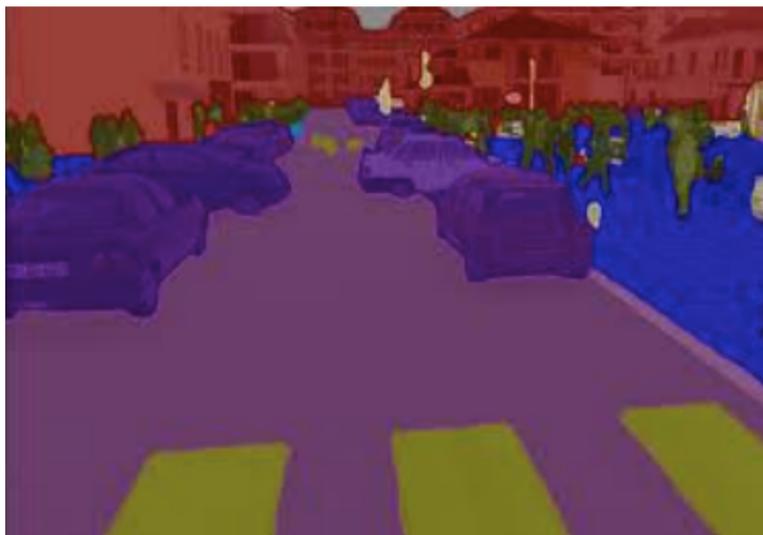
# Convolutional Neural Networks (CNN)

Le **CNN** si basano su strategie che diminuiscono la sensibilità alla posizione assoluta di elementi in un'immagine, aumentando la robustezza al noise

- zoo di architetture incredibilmente vasto, guidato dai progressi in applicazioni di computer vision
- versioni "deep" delle CNN possono estrarre caratteristiche molto complesse da immagini ad alta risoluzione
- translational-invariant feature learning → alta versatilità nell'applicazione di layer convoluzionali a problemi diversi

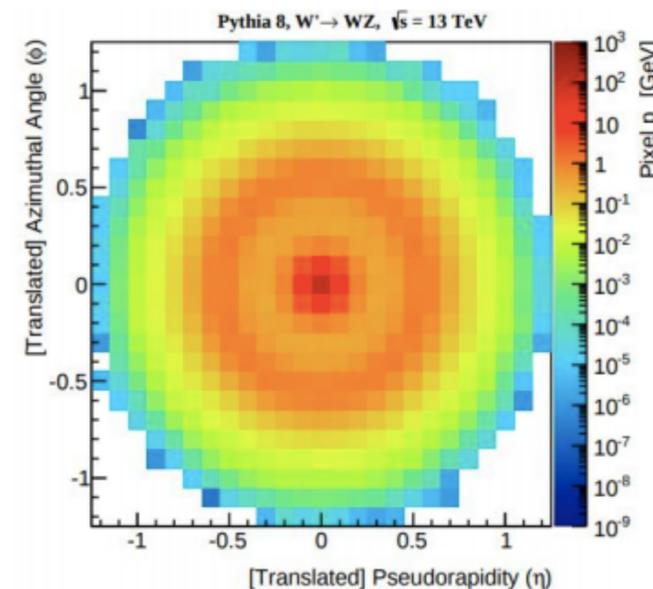
## Industry:

ampia adozione in applicazioni di computer vision (es, self-driving cars, ..)



## HEP:

3D imaging nei rivelatori, event classification, automazione di histo checking, etc



Tattica generale (TPCs, CALOs..): rappresentare i dati come immagini 2D/3d (anche 4D con timing info)

→ **problem casting in un task di computer vision**

# Deep learning in jet flavour identification

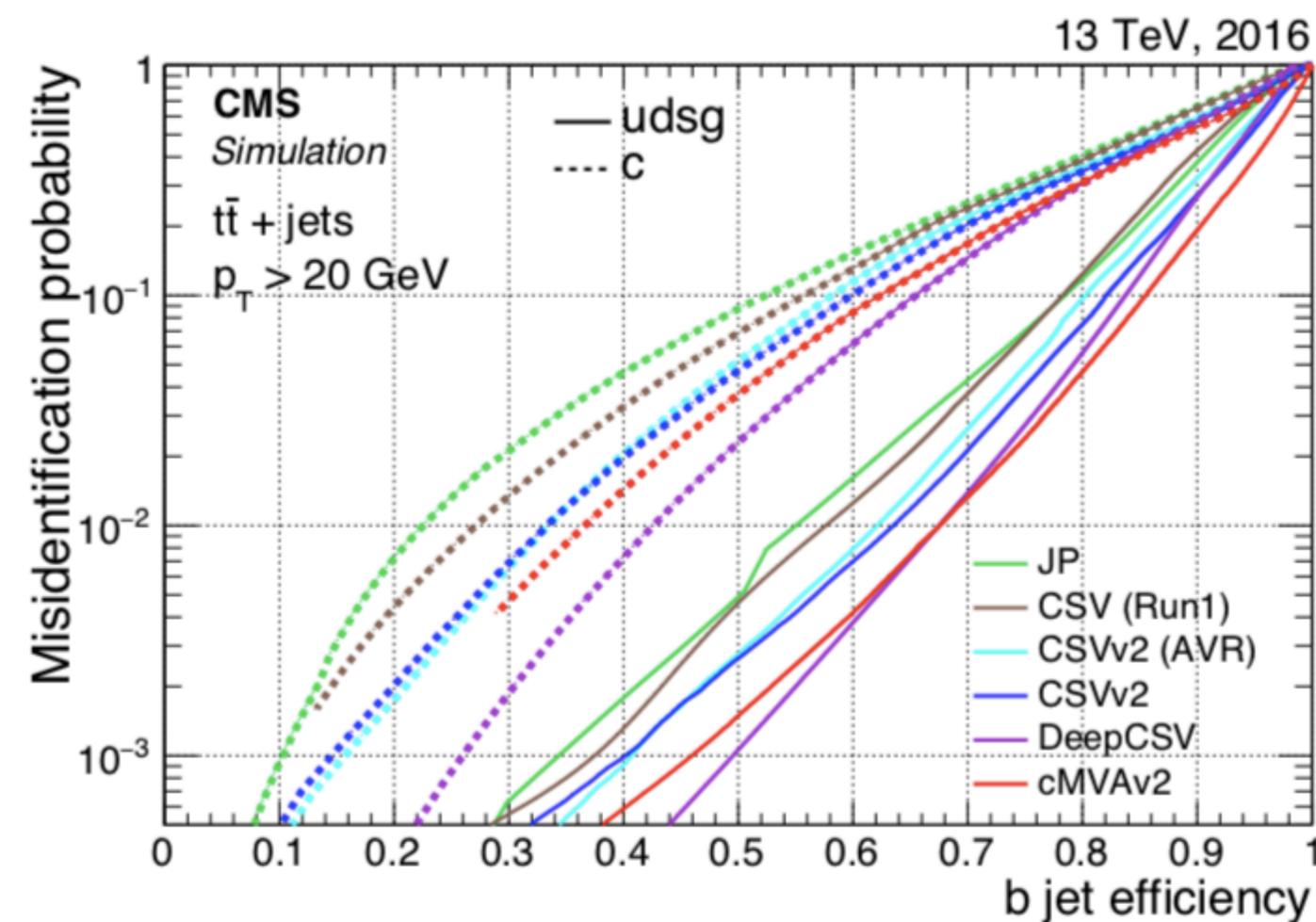
[1] JINST 13 (2018) P05011

[2] JINST 15 (2020) 12, P12012

Difficile rappresentazione high-level → alimento DeepNN con grandi moli di low-level features

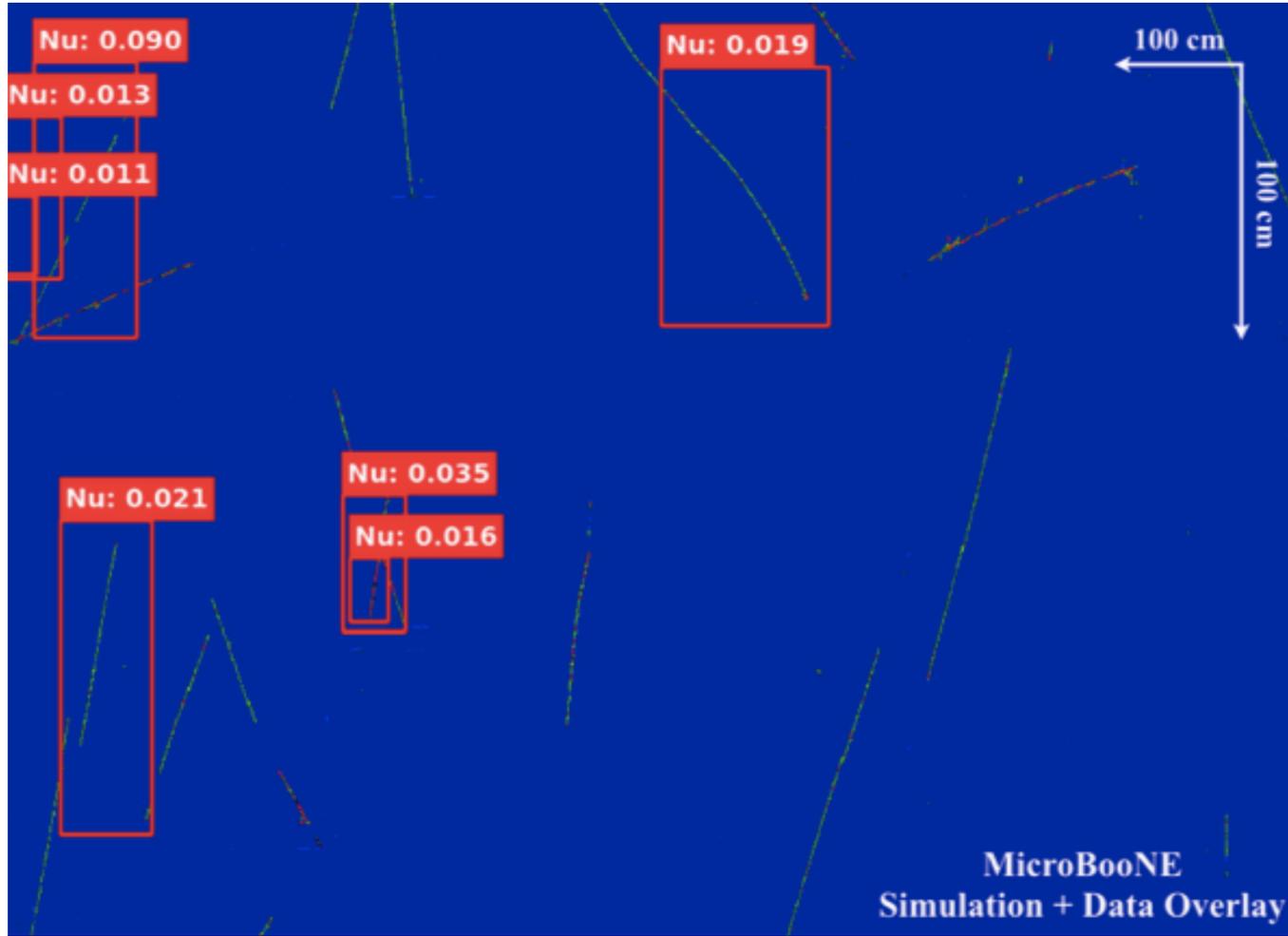
Es. identificazione di heavy-flavour (b/c quarks) jets in collisioni pp at 13 TeV di **CMS**

- DeepCSV algorithm [1]: versione con **NN** del combined secondary vertex (CSV) tagger usato in LHC Run-1  
→ **miglioramento relativo del 15% nell'efficienza** relativa w.r.t metodi likelihood-based
- DeepJet [2]: usa lista completa di candidati particle flow, vertici secondary, .. → **ulteriori miglioramenti**, ed estensione a quark-gluon tagging



Esempio

# "HEP is so different from other applications.. no?"



Rivelazione di tracce di **neutrini** su cosmic background events (metodo: **CNN**)



Rivelazione di **aeroporti** da immagini satellitari (metodo: **CNN**)

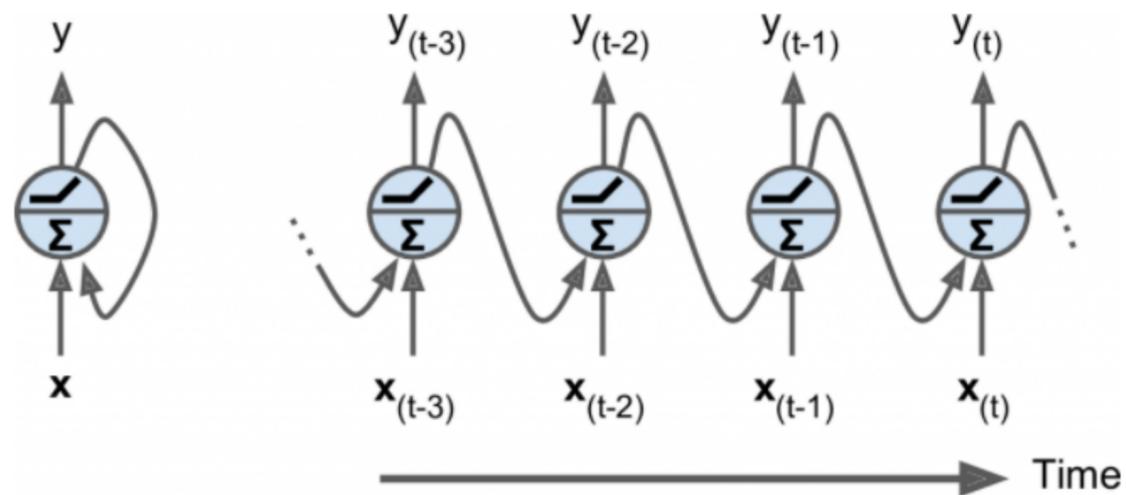
# Recurrent Neural Networks (RNN)

Le **RNN** si basano su "recurrent neurons" (backward-pointing connections) e permettono di trattare input di lunghezza variabile e processare time-series accumulando e usando insieme info da istanti diversi nella sequenza

- applicazioni in serie temporali di ogni genere, es. language translation ..

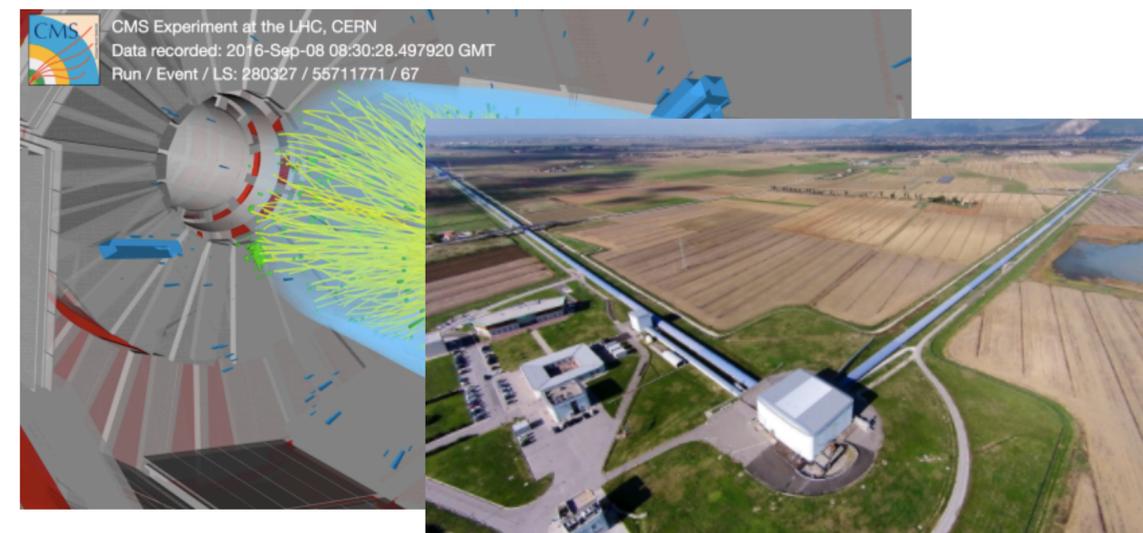
## Industry:

gestione di "time series"  
(audio, video, natural language processing)



## HEP:

classificatori capaci di processare segnali complessi e input di natura variabile (tracks, particles in jets, etc) - ampie applicazioni correnti e future in astro-particle physics



# RNN per l'identificazione di b-quark jets

[1] ATL-PHYS- PUB-2017-003 and 013

[2] CMS-DP-2017-005

[3] Phys. Rev. D 94, 112002 (2016)

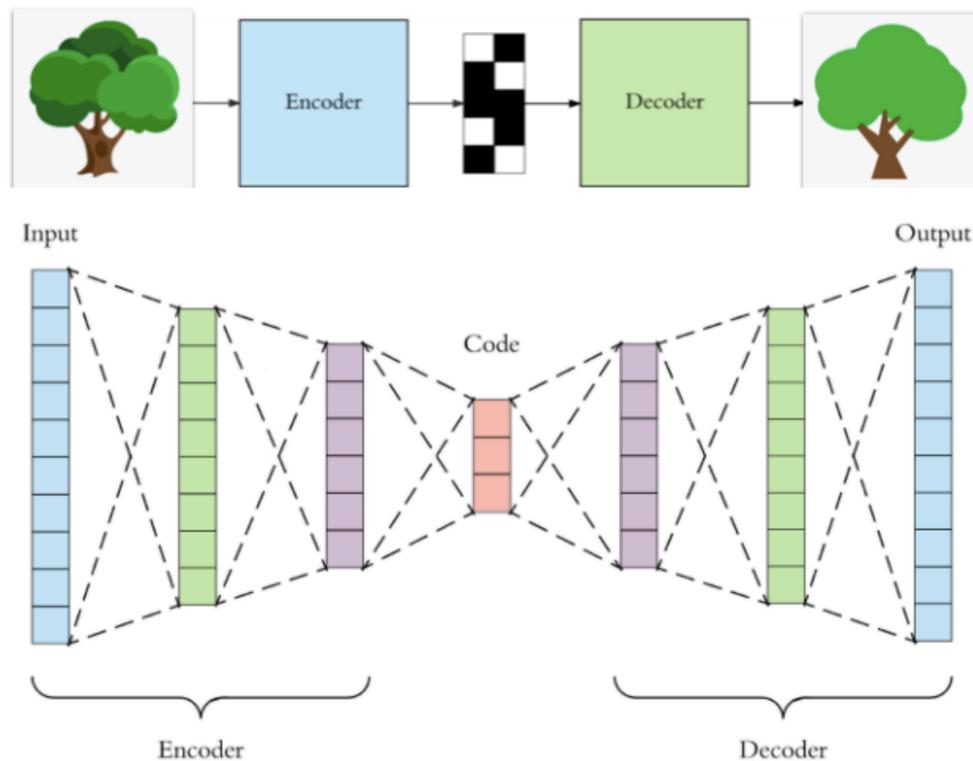
b-tagging con **RNN** può usare low-level features di particelle nei jet

- ordinamento delle jet particles in una sequenza, ranking sull'incompatibilità con il primary vertex, estrazione di un set di feature per ogni particella, input alla RNN, allenamento per discriminare tra b-quark jets e altri jets
- → **ATLAS mis-identification rate ridotto di ~4x** rispetto ad algo non-ML con un'**aggiuntiva riduzione ~3x** quando la RNN viene usata come input feature nel successivo training di BDT o NN [1]. Simili risultati da **CMS** [2] e promettenti sviluppi con strutture RNN più sofisticate [3]

# Autoencoder (AE)

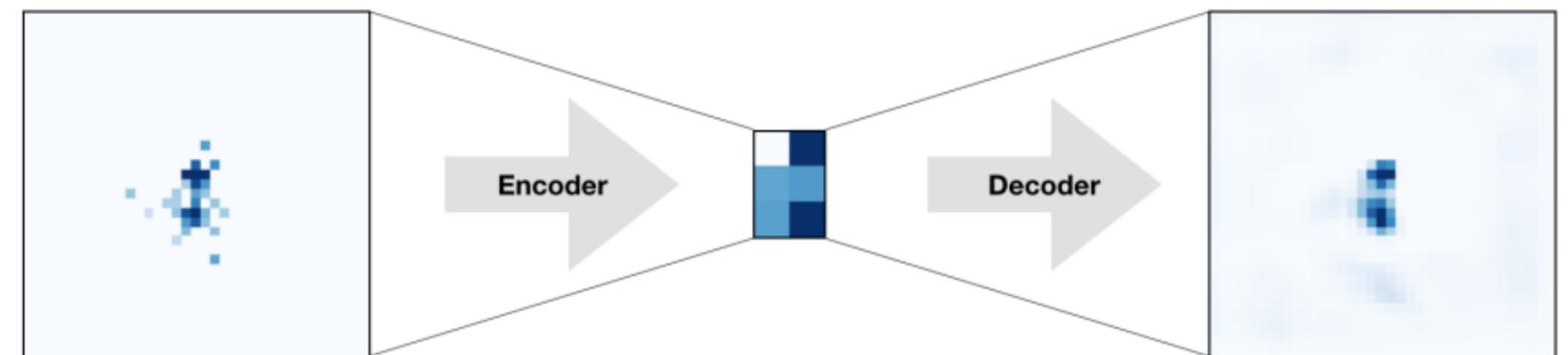
**AE** è un algoritmo di compressione "data-specific". Come architettura è una NN feed-forward (unsupervised) in grado di comprimere l'input in una rappresentazione a dimensionalità ridotta ("latent space") e poi decomprimerla in output

**AEs in Industry:**  
dimensionality reduction (like PCA),  
clustering, denoising, ...



**AEs in HEP:**

anomaly detection (eventi interessanti sono quelli la cui decompressione in output è distante dall'input seconda una data metrica)



Potenziale strumento per scoprire nuova fisica in modo "unsupervised"

Un **Variational AE** è stato proposto come architettura da allenare su processi fisici noti e costruire delle "thresholds" per identificare eventi "anomali" come outliers, e costruirne un catalogo da scrutinare successivamente

- training indipendente da specifiche signature di nuova fisica → assumptions-free
- complementare a metodi classici (tipicamente basati su model-dependent hypothesis testing)
- topologie ricorrenti nel catalogo possono ispirare search focalizzate, e anche costruzione di modelli teorici
- inserire questa tecnica a livello di trigger (es. LHC) potrebbe rafforzarne efficienza e dargli versatilità

Un'idea per **unsupervised learning** in HEP basata su un'architettura di NN esistente, che potrebbe estendere la physics reach di esperimenti HEP

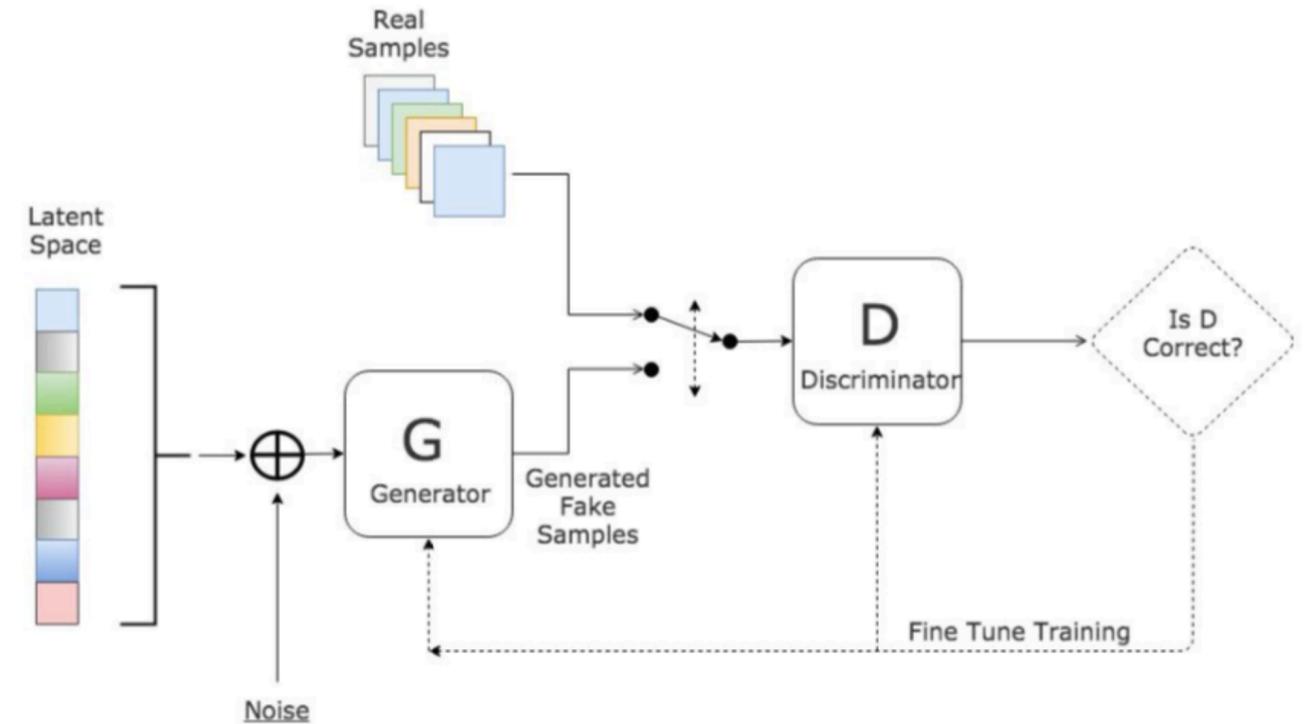
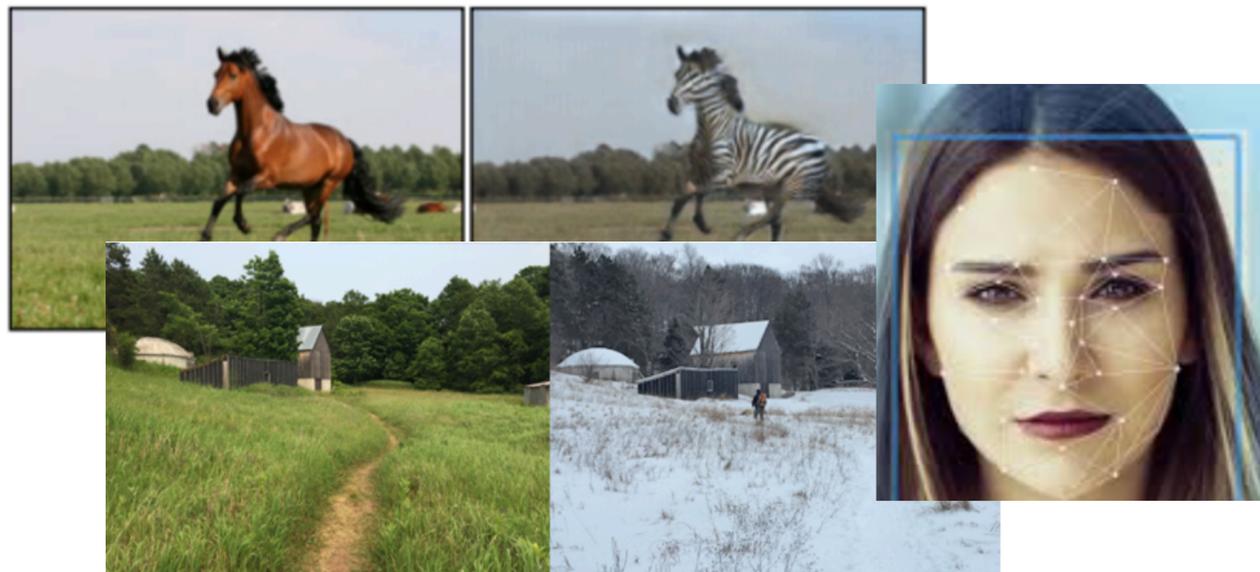
# Generative Adversarial Networks (GAN)

Un algoritmo generativo. L'architettura è una doppia NN di cui una (generatore G) crea immagini da rumore, e l'altra (discriminatore D) classifica le immagini di G come real vs fake

- allenate una contro l'altra, obiettivo di G è confondere D, imparando così a creare immagini fake ma realistiche

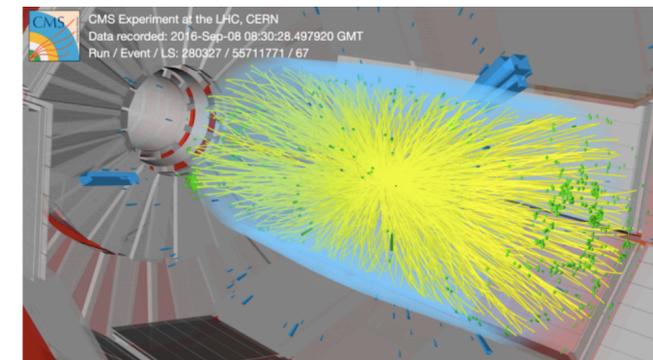
## Industry:

image editing, data generation, security, ..



## HEP:

Simulare la risposta del detector a costo computazionale molto ridotto



Esempio

# Event simulation con GAN

[1] Phys.Rev.D 97 (2018) 1, 014021

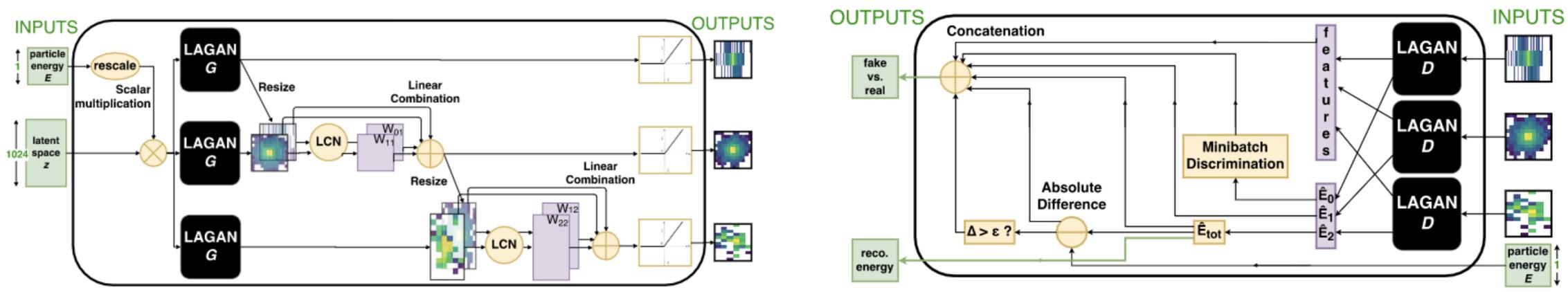
Modelli generativi veloci e ad alta fedeltà possono rappresentare alternative

- apprendere da data sample esistenti → high dimensional feature distributions sampling

Primi approcci mostrano miglioramenti in velocità delle simulazioni di ordini di grandezza rispetto a FastSim, di cui tutti gli esperimenti HEP beneficerebbero...

- → es. **CaloGAN**, una tecnica di FastSim per simulare 3D showers in sistemi ECAL multi-strato con delle GAN, può apprendere e generare l'immagine calo ricostruita saltando gli step GEANT e RECO → **10k faster than the default**

.. ma **l'accuratezza non è ancora sufficiente** (limitazioni del metodo, instabilità nel training, ..)



CaloGAN composite generator (left) and discriminator (right)

# Data sparsity e point clouds

HEP maneggia dataset ad **alta sparsity**: un problema comune a molti domini scientifici

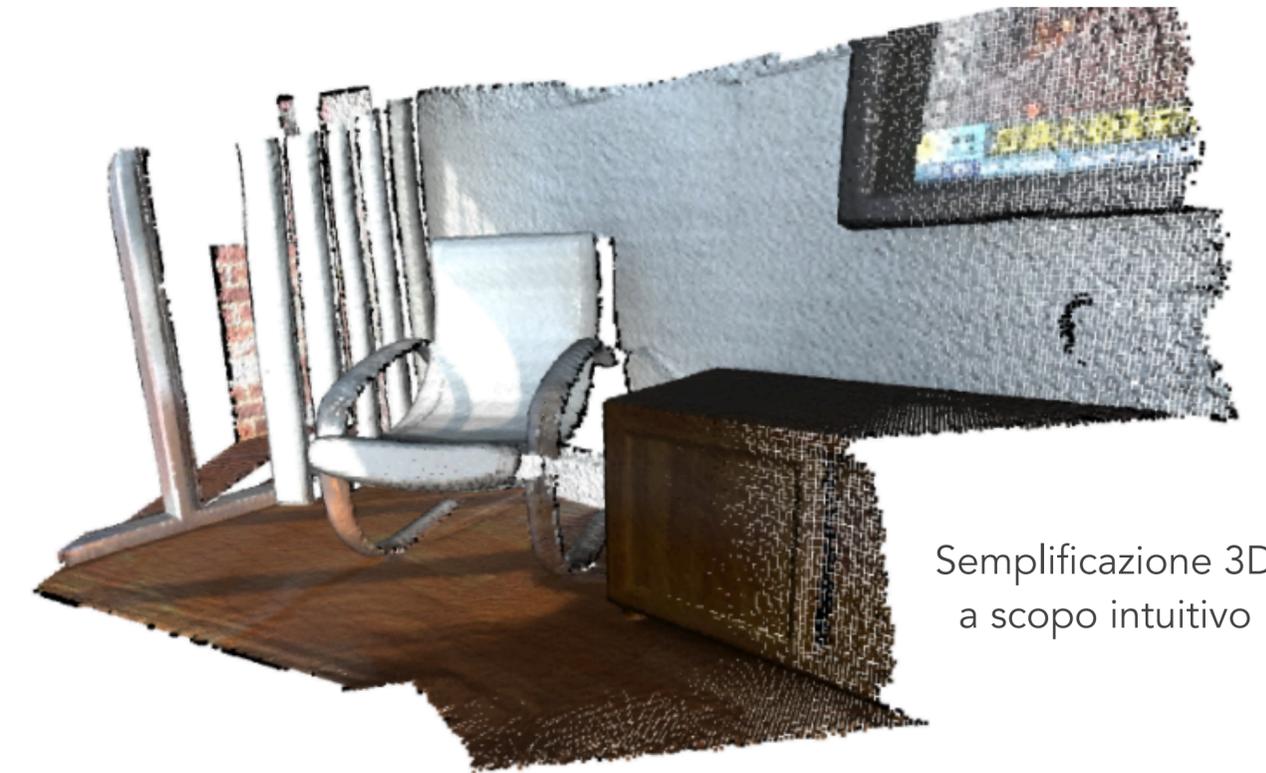
- granularità e occupancy dei sensori nei rivelatori HEP
- popolazione di stelle e galassie nel cosmo
- descrizione molecolare in chimica computazionale

Spazio astratto con coordinate di elementi sparsi, ciascuno caratterizzato da un array di features, vari arrays in funzione dell'evento/run, ..

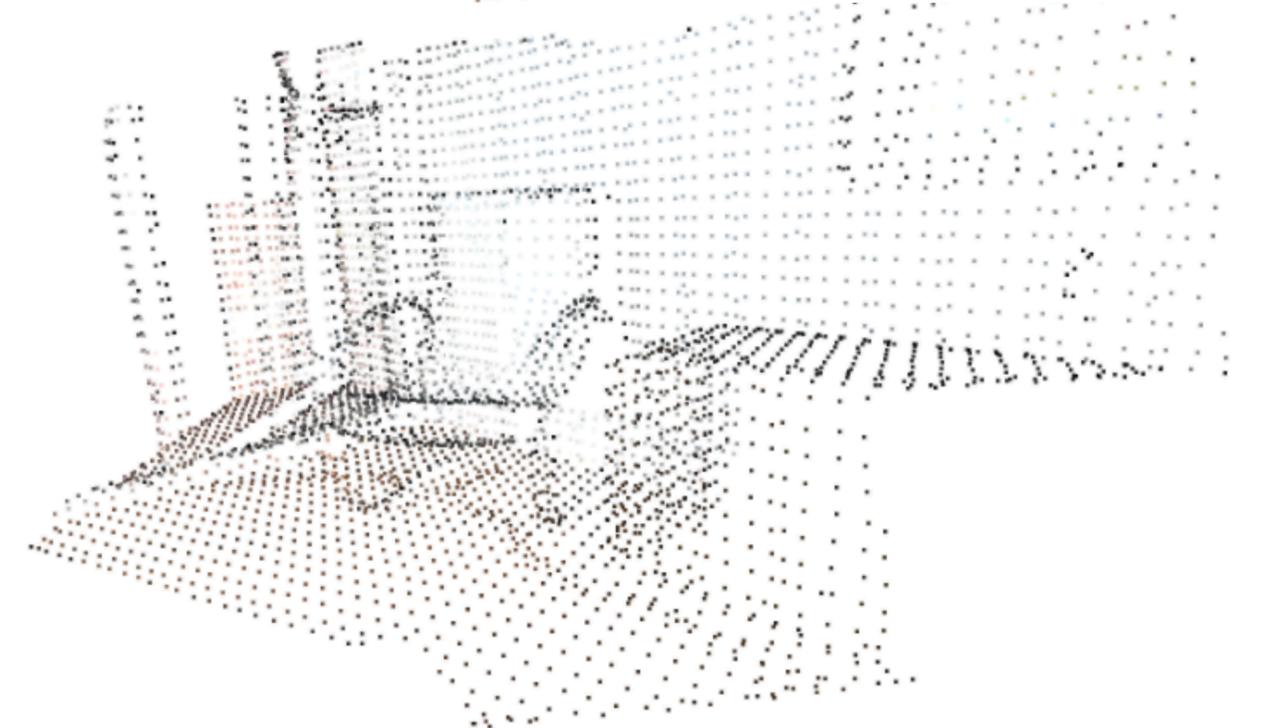
- e.g. sciame EM → energia depositata in volumi attivi di un ECAL

Rappresentazione adeguata è quella di una **"point cloud"** e l'approccio migliore potrebbe essere non tradizionale.. e richiede una certa apertura mentale verso metodi a noi non familiari..

- problema configurabile in modo sorprendentemente simile alla trattazione dei social media dataset..



Semplificazione 3D  
a scopo intuitivo



# Graph Neural Network (GNN)

La potenza di una CNN per le immagini si poggia (anche) sulla rappresentazione come array delle informazioni dense in esse contenute, con metrica intrinseca

- “regular-array dataset” paradigm

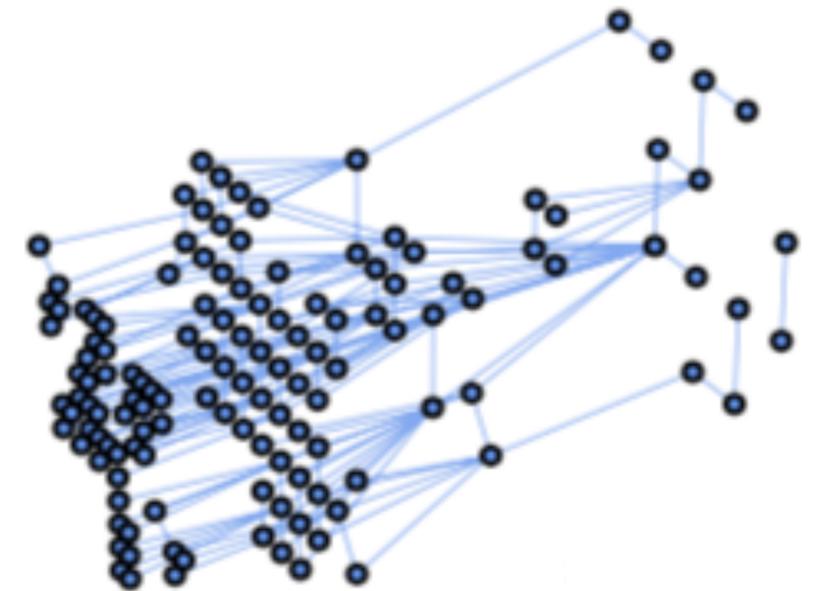
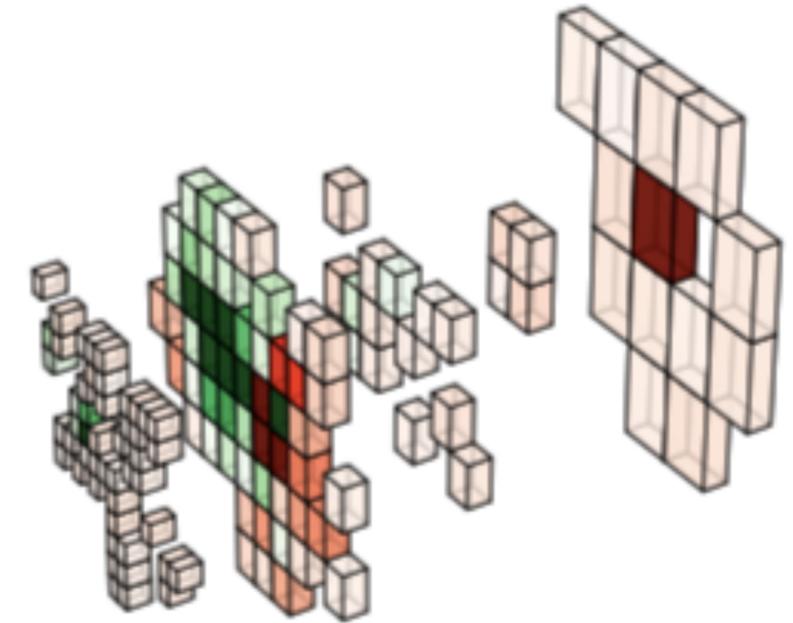
In una rappresentazione sparsa, serve invece una metrica che definisca la prossimità nello spazio astratto delle nostre features

Come? Passando da “**datasets**” a “**graphs**”

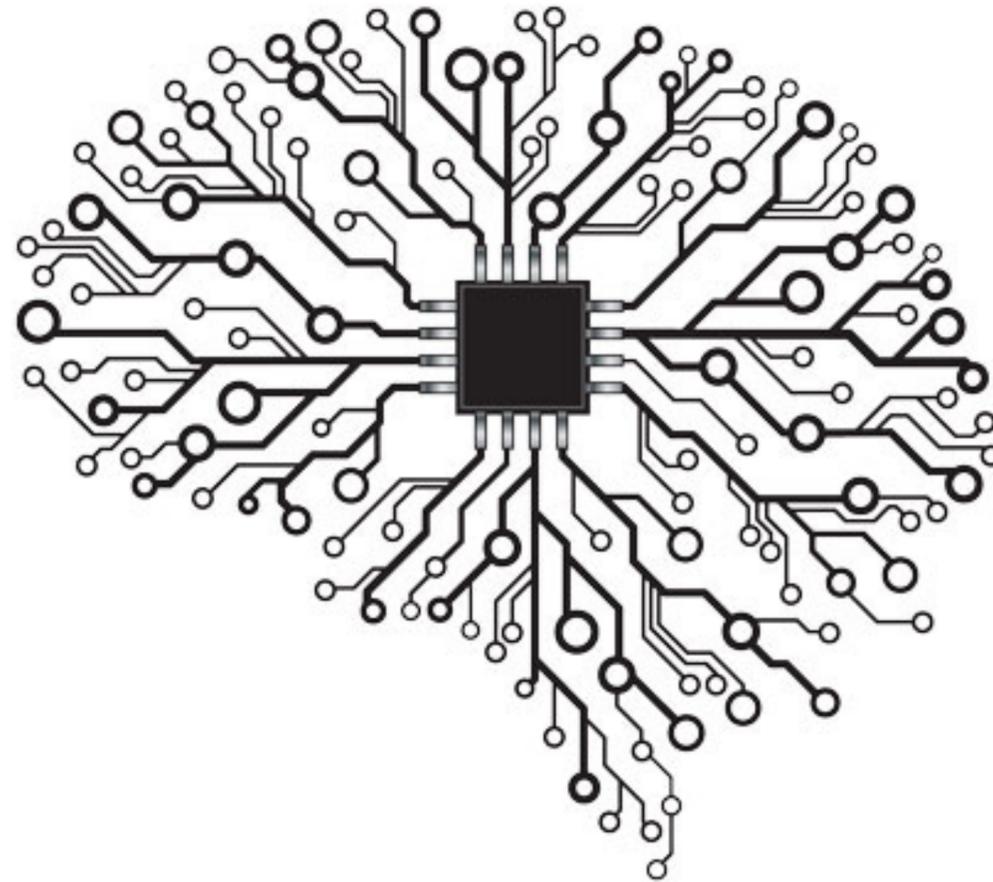
- connettere elementi di un dataset e allenare una NN a imparare quali sono le connessioni rilevanti

GNN → costruire una data structure,  $(V,E)$  con  $V$ =vertex e  $E$ =Edges, scelta possibile di vari tipi di vertici, senza un prior si fa un fully-connected graph, etc..

Le GNN potrebbero essere il (o parte del) futuro del DL in HEP

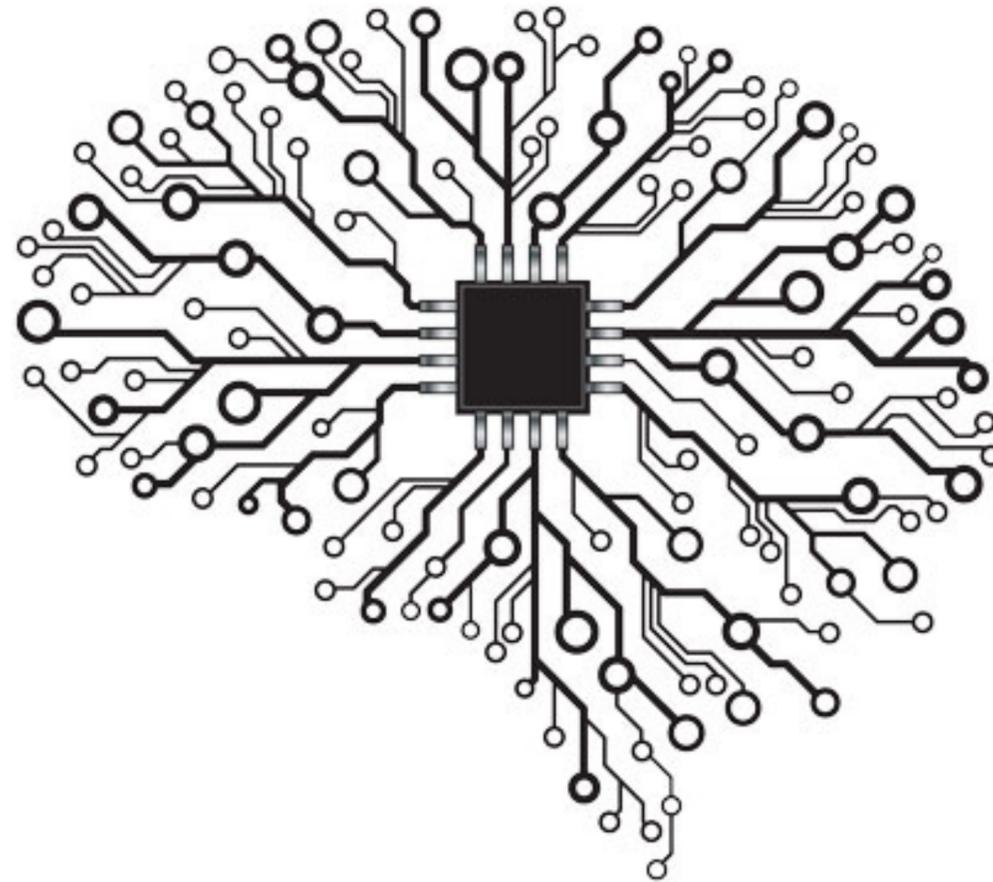


# Direzioni dell'AI → quali sono rilevanti per INFN?



# Direzioni dell'AI → quali sono rilevanti per INFN?

## Ubiquity

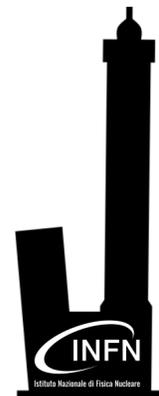
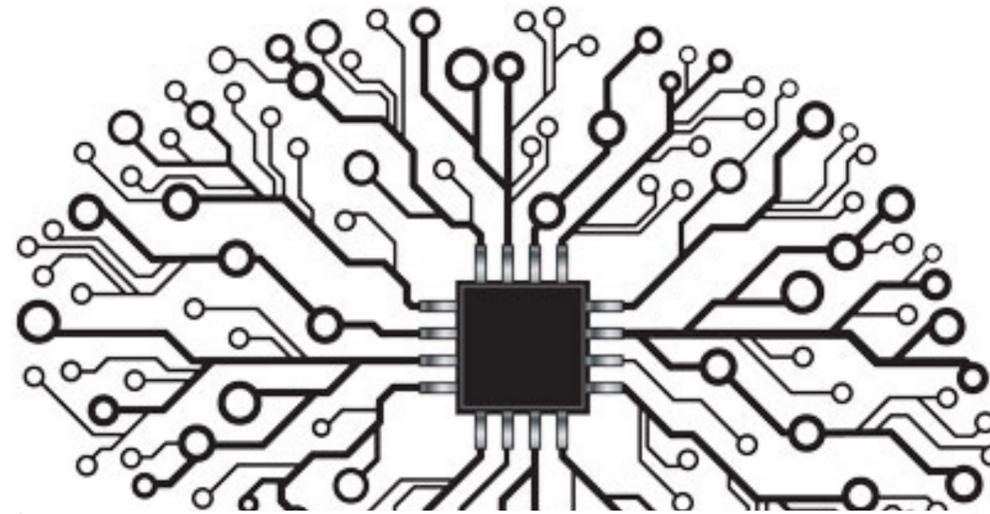


# Direzioni dell'AI → quali sono rilevanti per INFN?

Ubiquity

## Human-centered AI

- ritenuto il "core" della "third AI wave"
- Stanford, Berkeley, MIT hanno stabilito i primi istituti su questo fronte



<https://centri.unibo.it/alma-ai/en>

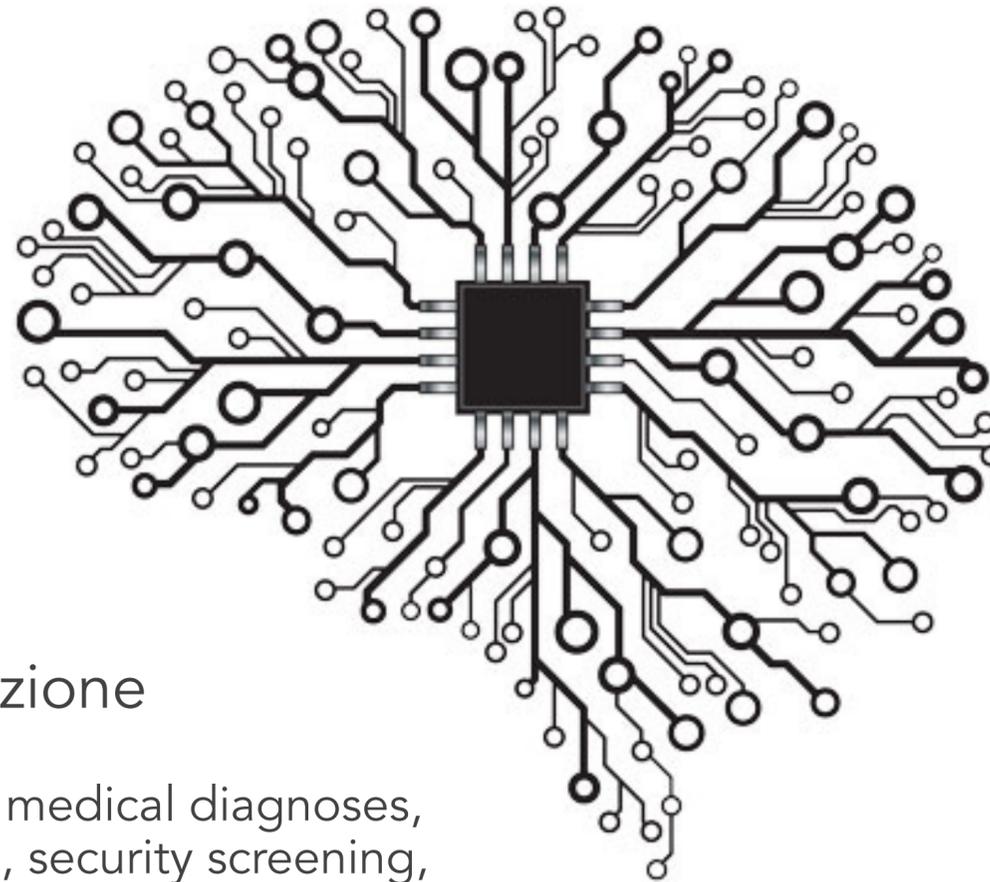
# Direzioni dell'AI → quali sono rilevanti per INFN?

Ubiquity

Human-centered AI

## Explainable AI (XAI)

- no "black box"
- aumentare fiducia e adozione
  - ❖ financial and legal decisions, medical diagnoses, industrial process monitoring, security screening, employment recruitment, legal judgment, university admissions, smart homes, autonomous vehicles, ..
- per INFN → interpretability, explainability, reproducibility of experiments



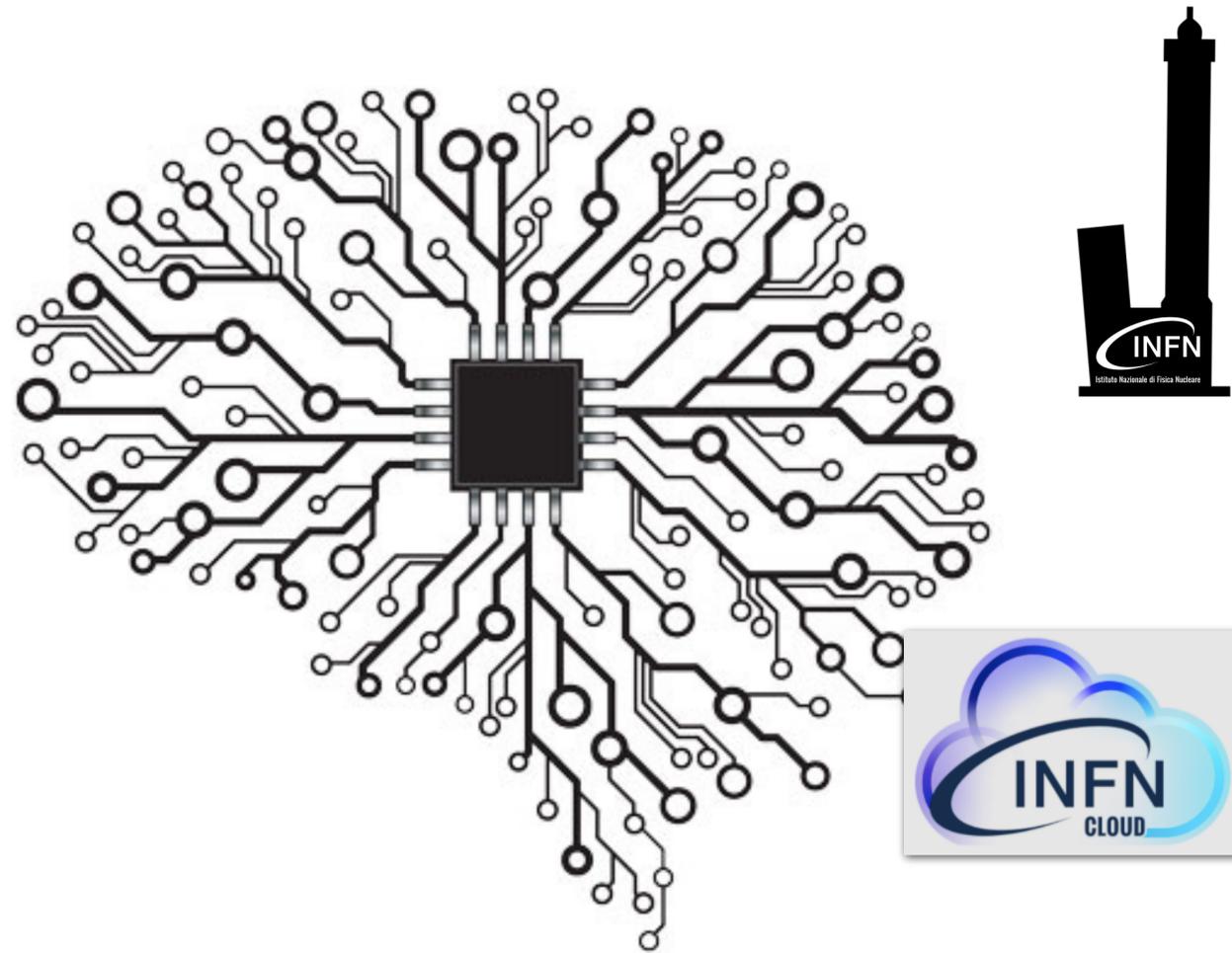
# Direzioni dell'AI → quali sono rilevanti per INFN?

Ubiquity

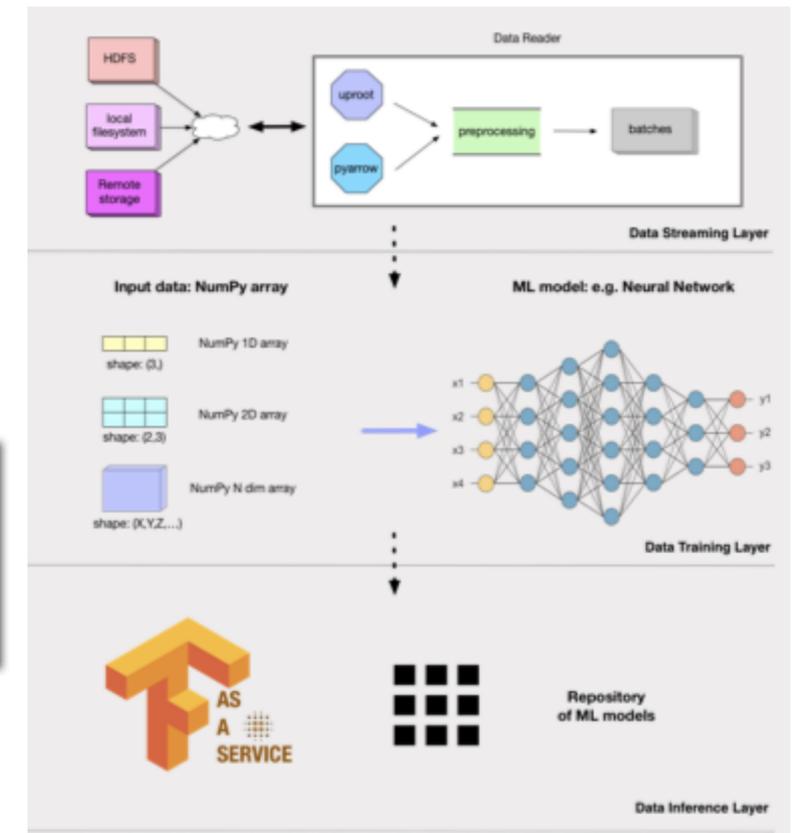
Human-centered AI

Explainable AI (XAI)

Useful and Usable AI



Attività di Bologna:  
ML as a service for HEP,  
coinvolgimento in ML\_INFN, etc



[Comp. Softw. Big Sci \(2019\) 5:17](#)

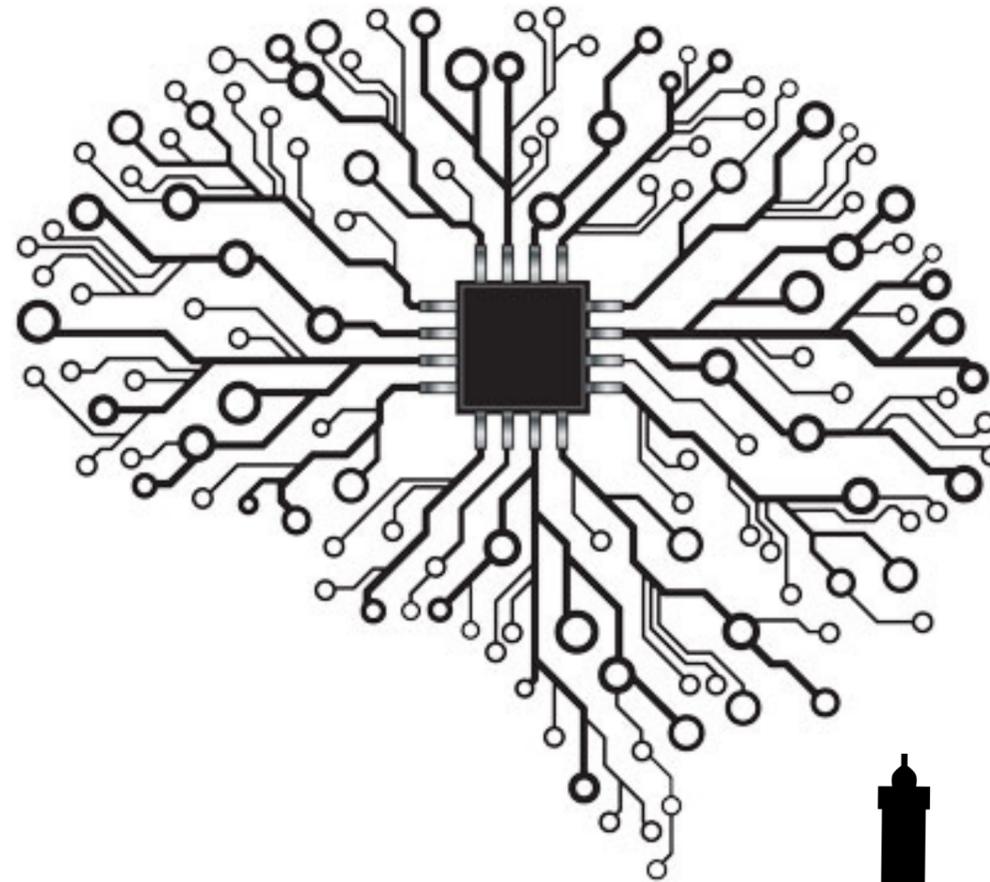
# Direzioni dell'AI → quali sono rilevanti per INFN?

Ubiquity

Useful and Usable AI

Human-centered AI

Explainable AI (XAI)

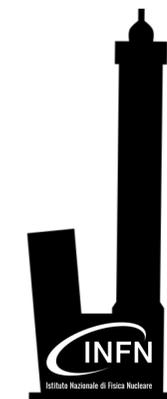


Green AI

- costo computazionale del DL dal 2012 al 2018 si stima sia cresciuto di un fattore 300,000x (!) → large carbon footprint
- NN non efficienti come il cervello, costi molto alti (tecnologie non inclusive), metriche di "energy efficiency"

**Attività di Bologna:**

Apertura attività sul fronte Green Computing (borsa PhD PON React, connessione con Ingegneria UniBO, etc)



# Direzioni dell'AI → quali sono rilevanti per INFN?

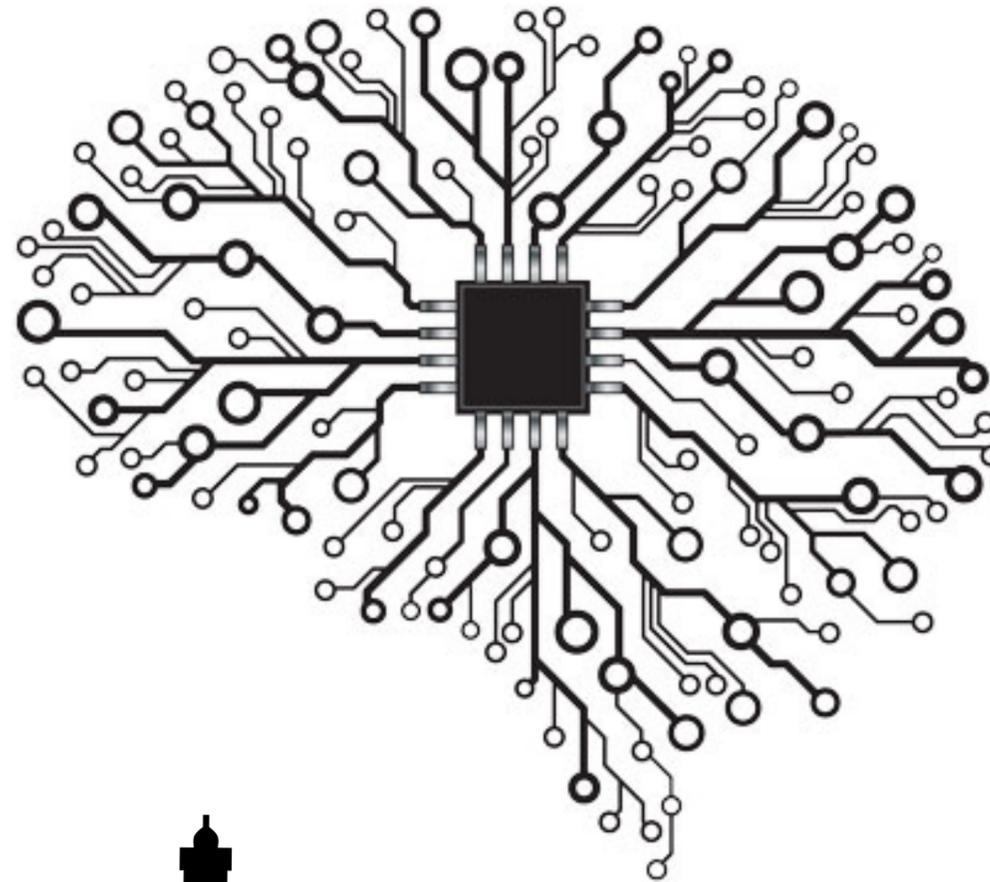
Ubiquity

Useful and Usable AI

Human-centered AI

Green AI

Explainable AI (XAI)

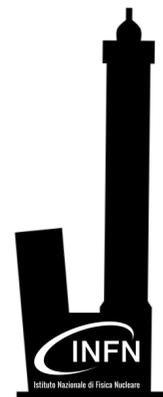


Quantum ML

- the near-term perspective on NISQ devices
- ❖ **circuit-based** QC → outsource della parte di predizione di un modello QML allenato classicamente
- ❖ **quantum annealers** → outsource della parte di training i.e. ottimizzare i modelli classici mappando un problema di ottimizzazione su un'istanza QUBO

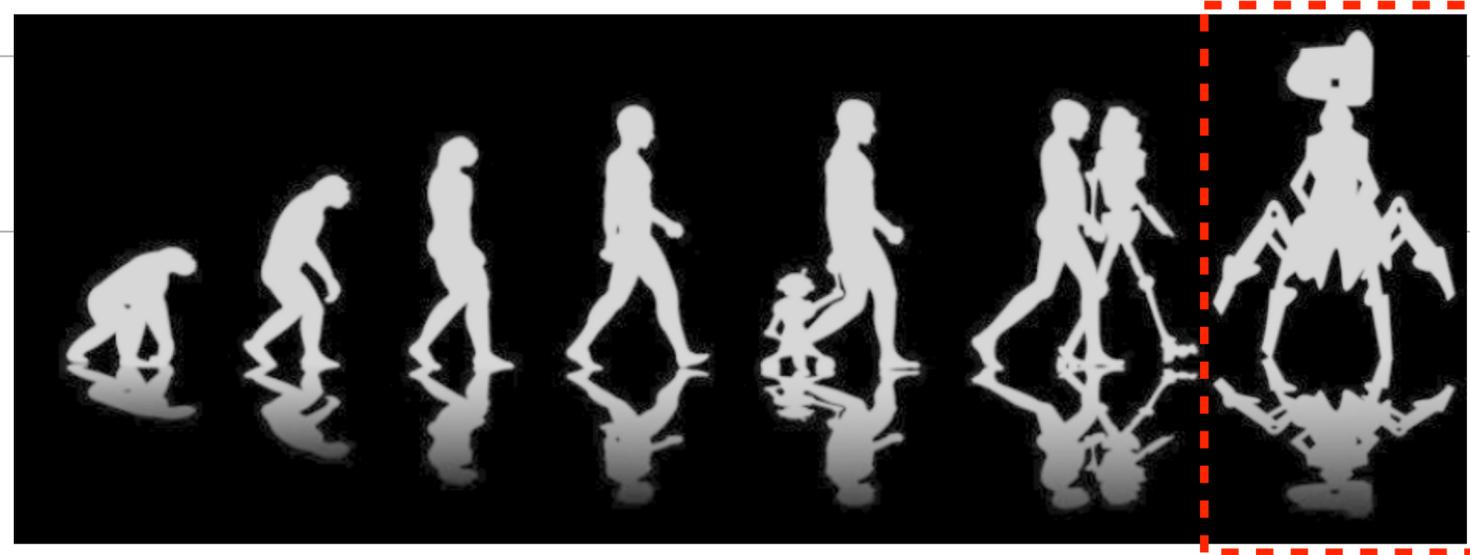
**Attività di Bologna:**

Apertura attività sul fronte Quantum, mappatura UniBO, connessione CERN-INFN-Univ, attività nuove in PNRR, ..



65<sup>^</sup> INFN-Bologna / D.Bonacorsi

# Summary



(AI/)**ML/DL** sono ubiqui in HEP, continuano a crescere e sono qui per restare

- le tecniche usate e i livelli di adozione variano molto a seconda dei contesti
- i metodi usati si evolvono molto più in fretta di quanto siamo in grado di usarli in HEP

Punti chiave:

- **cooperazione tra HEP e comunità ML** : uscire dalla nostra comunità, aprire linee di lavoro e progetti che portino interesse e competenze della comunità ML ai nostri (HEP) problemi, lavorare in vera interdisciplinarietà con altre discipline che hanno problemi comuni ai nostri
- investire sulla formazione di **ragazze e ragazzi giovani su skill in fisica computazione e/o metodi avanzati di software/computing**, e agire in modo concreto sul loro **riconoscimento di carriera**

# Thanks for the attention

## Credits

[1] Nature 560, 41–48 (2018)

[2] J. Phys. Conf. Series 1085 (2018) 022008

[3] Comp. Softw. Big Sci. 3, 7 (2019)

Thanks for material, discussion, inspirations to:

- M. Pierini, J.R. Vlimant, P. Vischia, M. Williams, S. May, S. Gleyzer, D. Bourilkov, co-authors of “*ML at the energy and intensity frontiers of particle physics*” [1], co-authors of the “*ML in HEP - Community White Paper*” [2], co-authors of “*A Roadmap for HEP Software and Computing R&D for the 2020s*” [3], coordinators of and contributor to the CMS ML forum, coordinators of and contributors to the IML team. With apologies to everyone I might have not mentioned.