deep ai image editor

Overview of Artificial Intelligence at RHIC and Beyond

Hannah Bossi (MIT)

RHIC/AGS Users Meeting 2024

Brookhaven National Lab

June 11th, 2024

MIT HIG's work was supported by US DOE-NP

Massachusetts Institute of Technology

Roadmap

What is AI/ML and why is it useful for physics?

How is AI/ML currently being used at RHIC?

What are some uses of AI/ML beyond RHIC?

- Note: This will be a brief and biased overview that aims to provide context for the remainder of the talks in this workshop!
  - Plus some commentary on ongoing challenges.
  - See the rest of the talks today for more specific details!

Roadmap

What is AI/ML and why is it useful for physics?

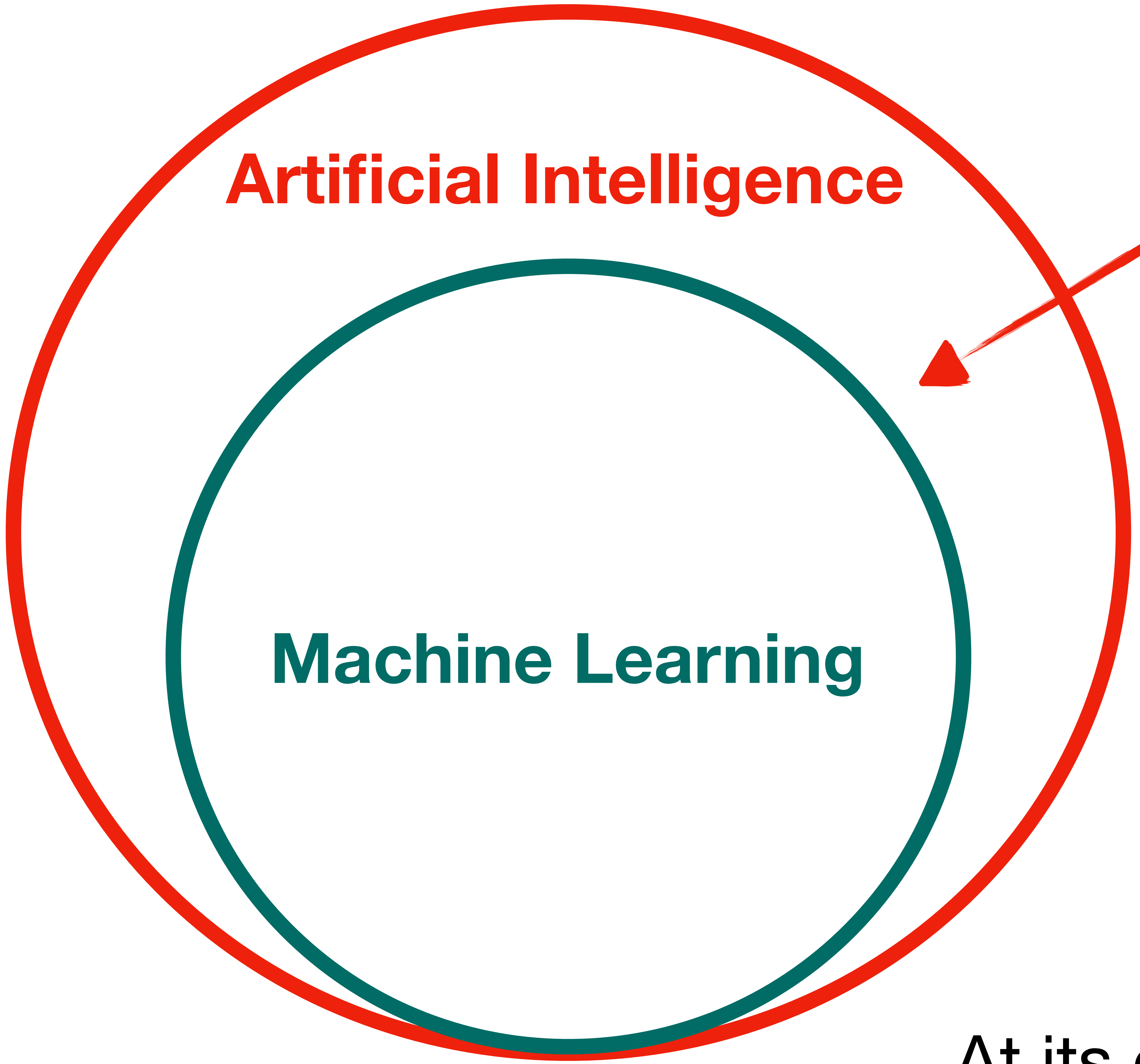
How is AI/ML currently being used at RHIC?

What are some uses of AI/ML beyond RHIC?

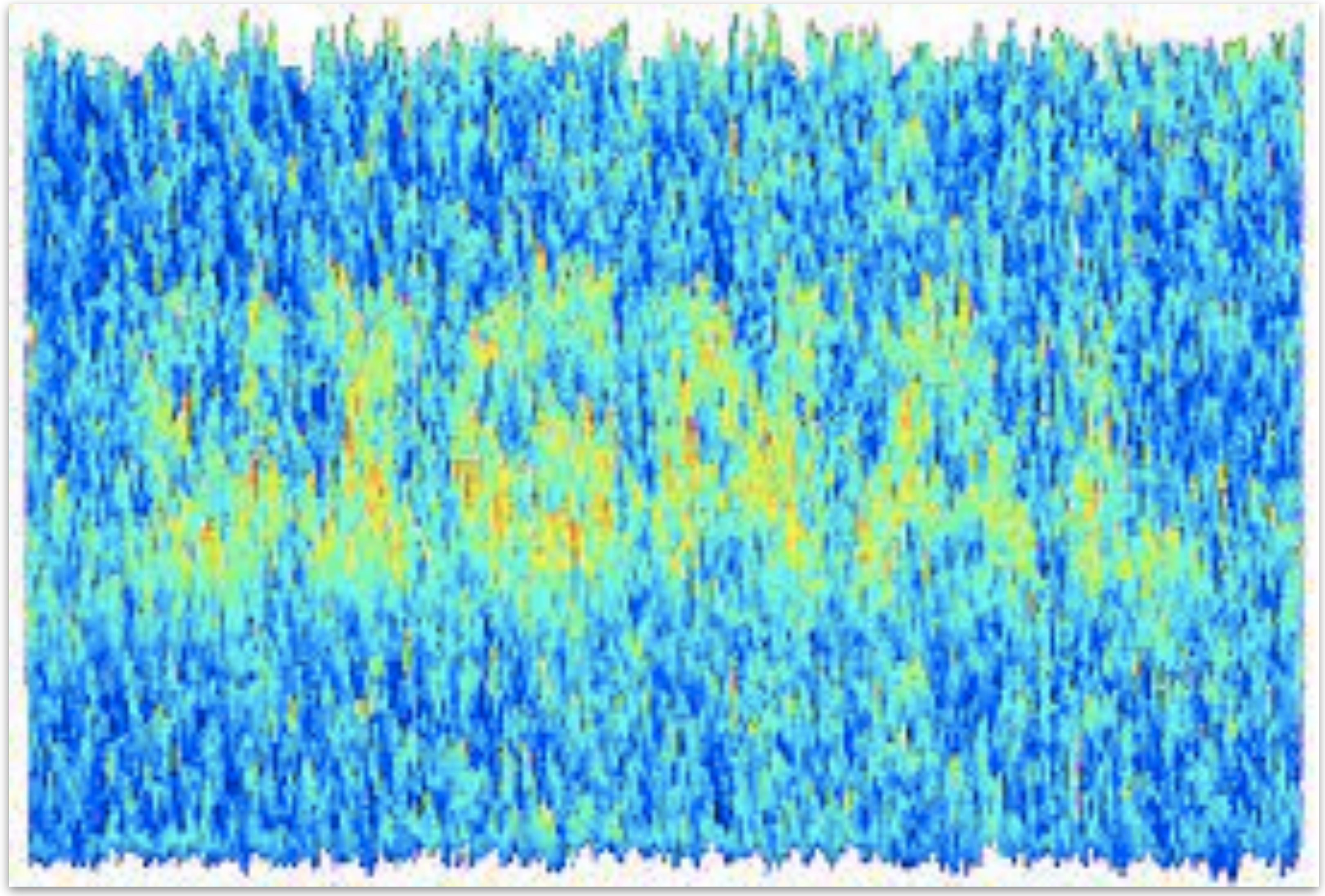
What is AL/ML?

Artificial Intelligence: Programs with the ability to acquire and apply knowledge and skills.

Ex: Chatbots (humans give rules)



Machine Learning: algorithms that imitate human learning, i.e. gradually improving accuracy over time.

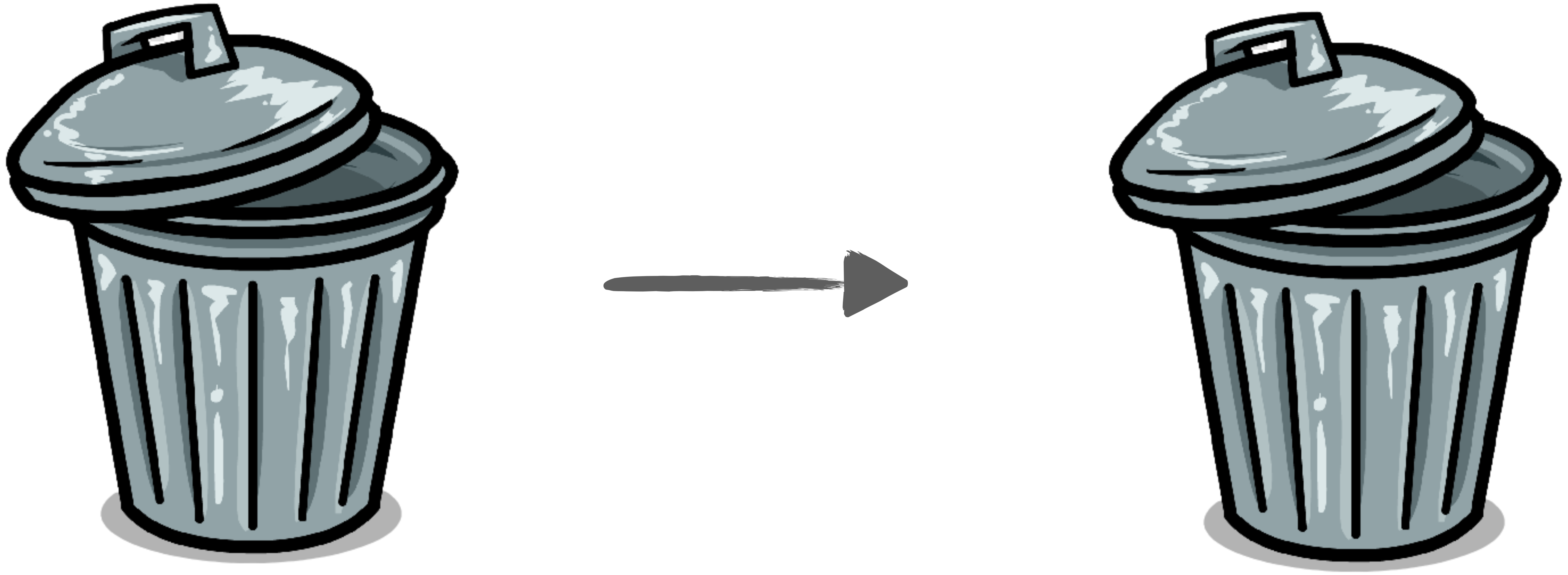


At its core, pattern recognition → humans can do this by eye!

What kinds of ML tools are there?

Best tool depends on the problem! (More details on actual algorithms in backup)

What can ML *not* do?



Garbage In *Garbage Out*

- ML cannot replace domain knowledge.
- ML is not a causation tool.

Don’t want to be finding cloudy days when you should be finding tanks!

ML is not a magic fix!

Why ML and physics?

Goal of experimental measurements: *To extract physics information from available data!*

Conventional approach: (1) make selection using a series of boolean decisions motivated by physics/experimental constraints (2) perform a statistical analysis on selected data.

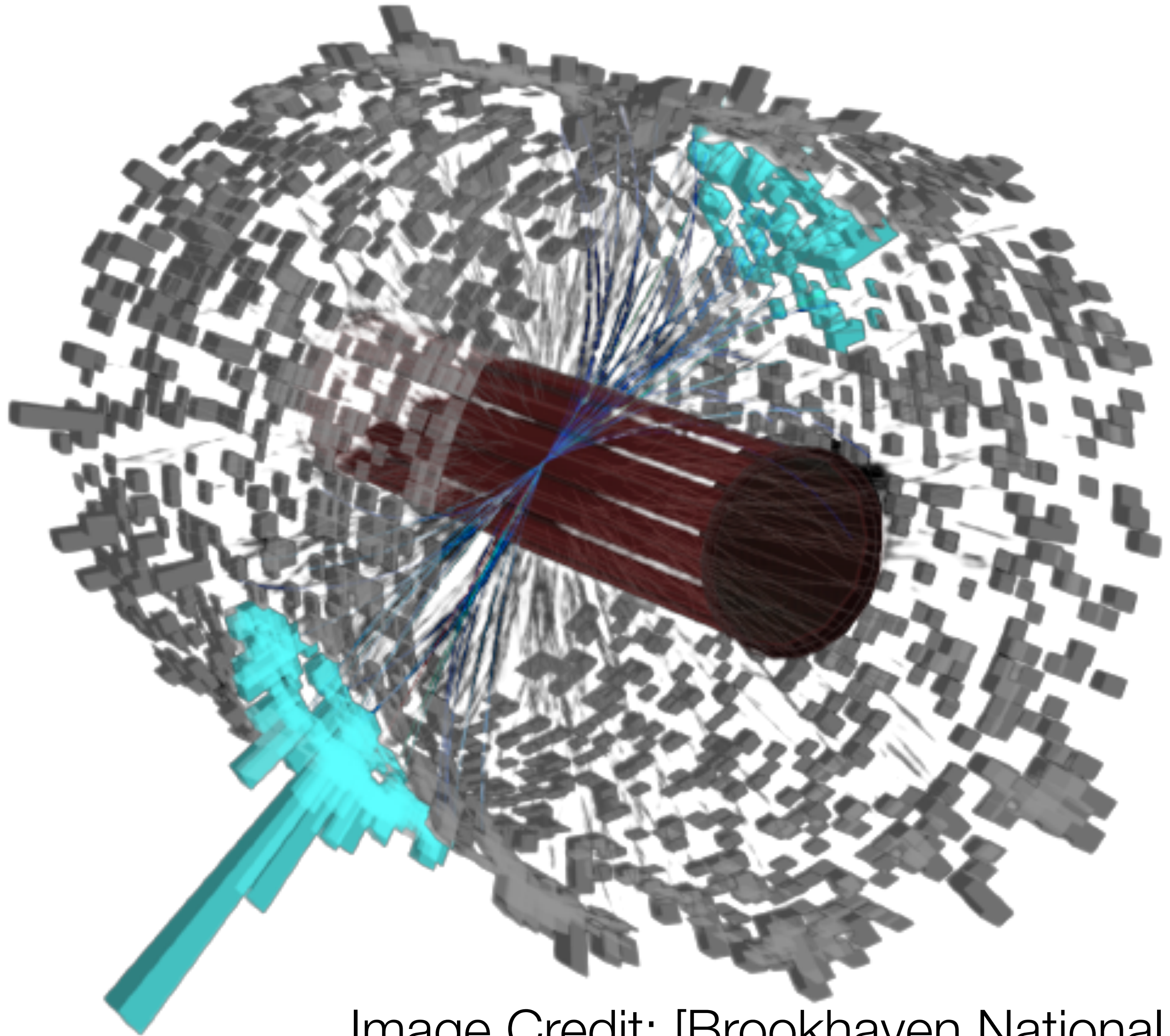
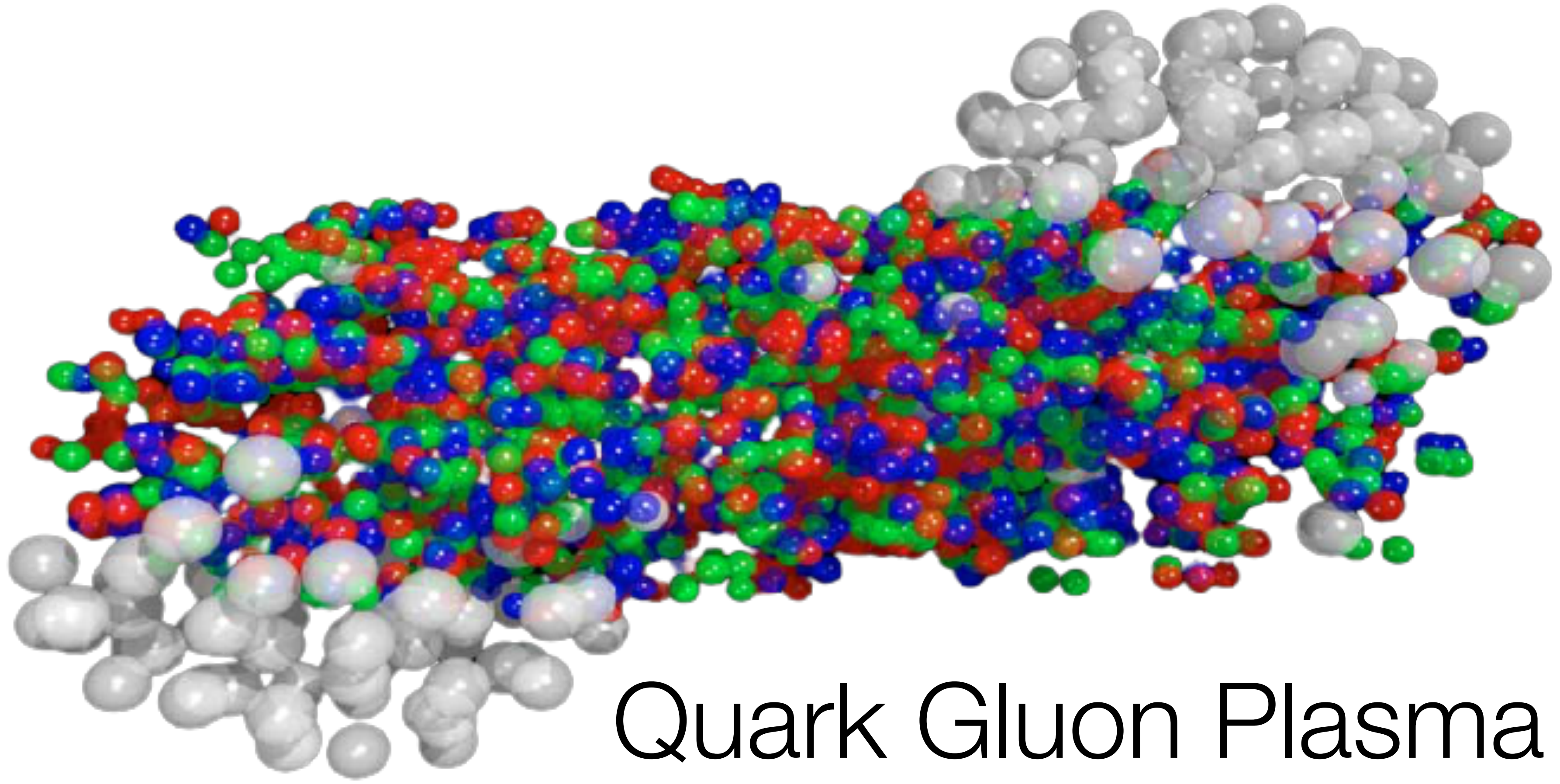


Image Credit: [Brookhaven National Lab]



Quark Gluon Plasma

Optimal decision difficult to derive from expert knowledge alone! Employ algorithms that utilize multiple variables simultaneously → inspired countless ML applications! [Living Review]

Big data @ BNL

Industry Academia

- SDCC passed 100 PB of stored data in 2017.
- **As of yesterday, total size was ~285 PB**
- sPHENIX alone is ~19 PB
- Dramatic increase compared to 1.2 PB in 2001!

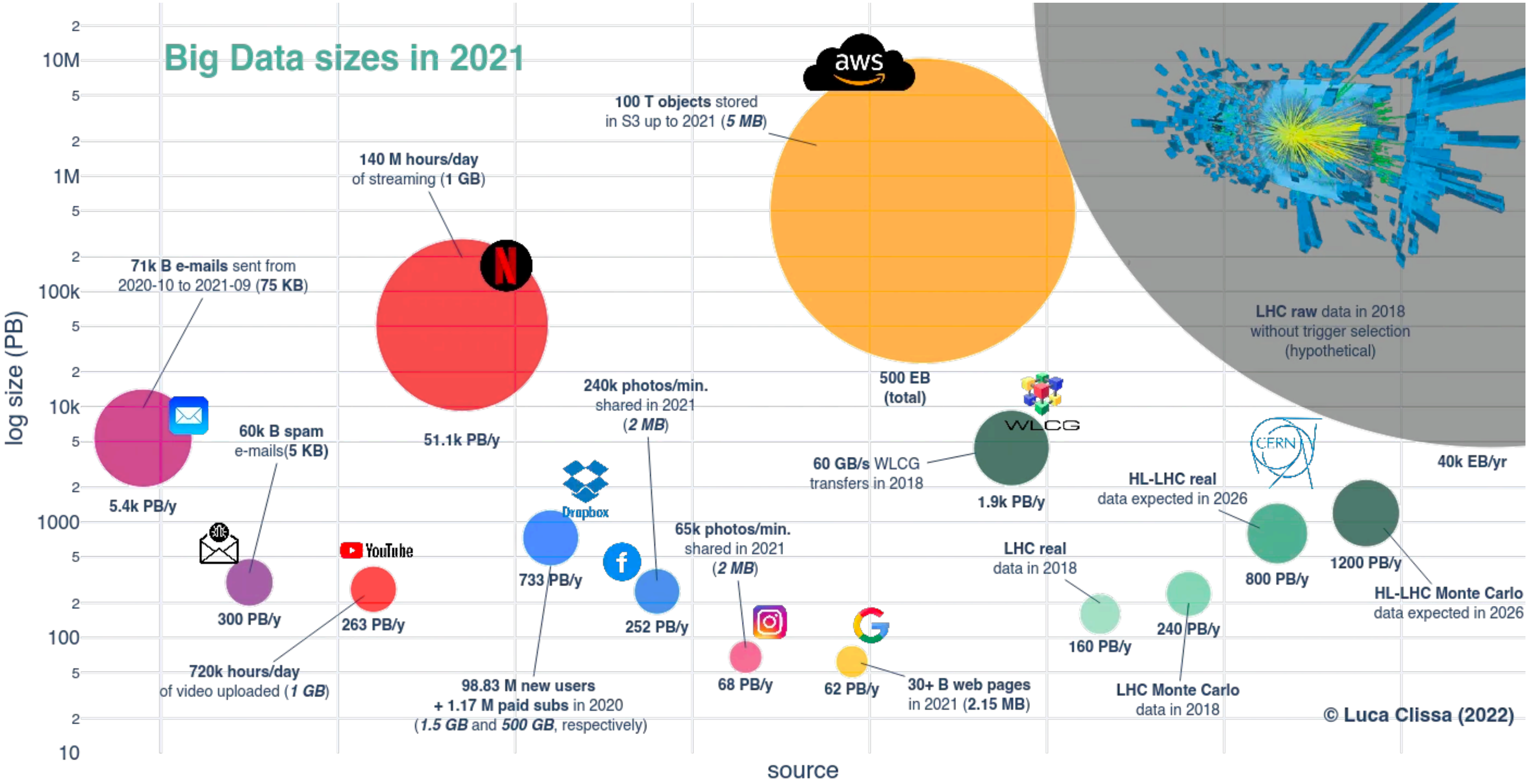
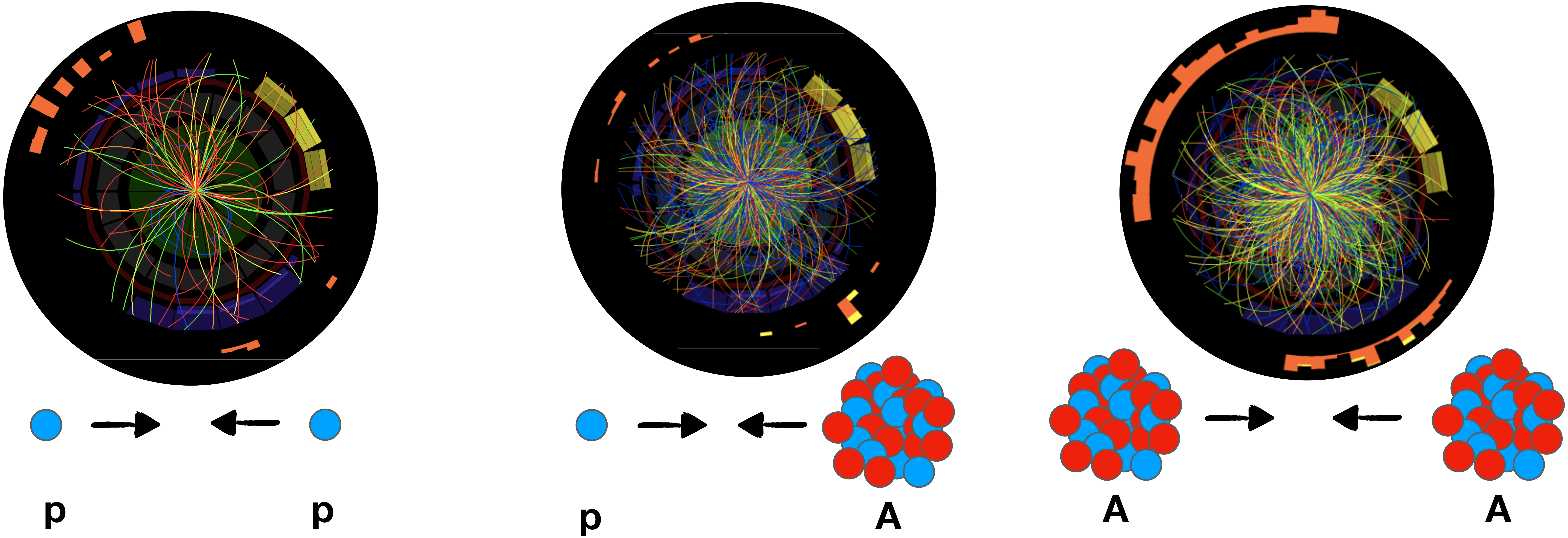


Image Credit: [Towards Data Science]

Data volumes comparable to medium-sized industry applications.

ML and heavy-ion physics

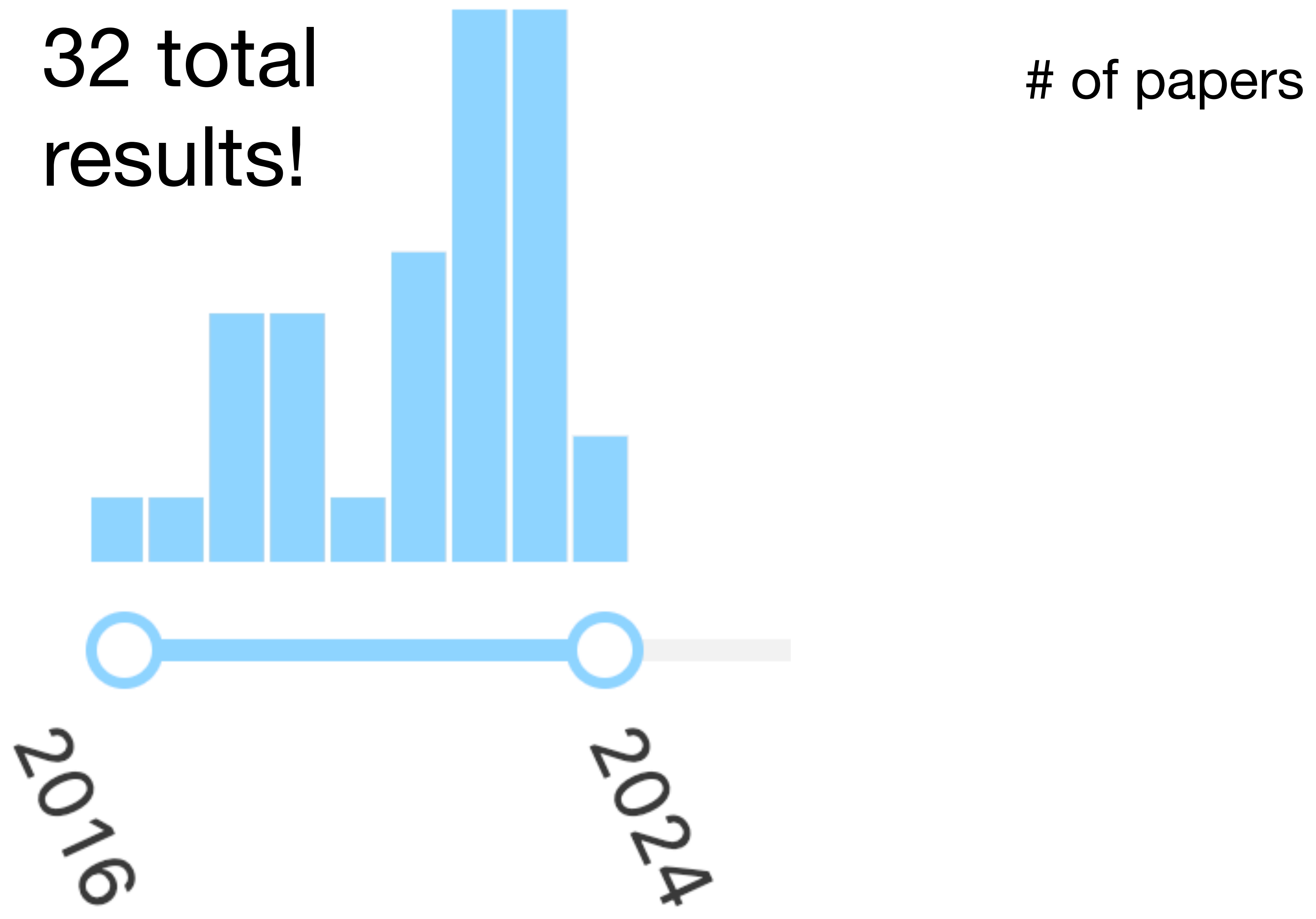


HI environment can be challenging for ML.

- Higher particle multiplicities, much more complex system (even by eye)!
- Training difficult due to dependence on simulation used in training

ML at RHIC and the EIC

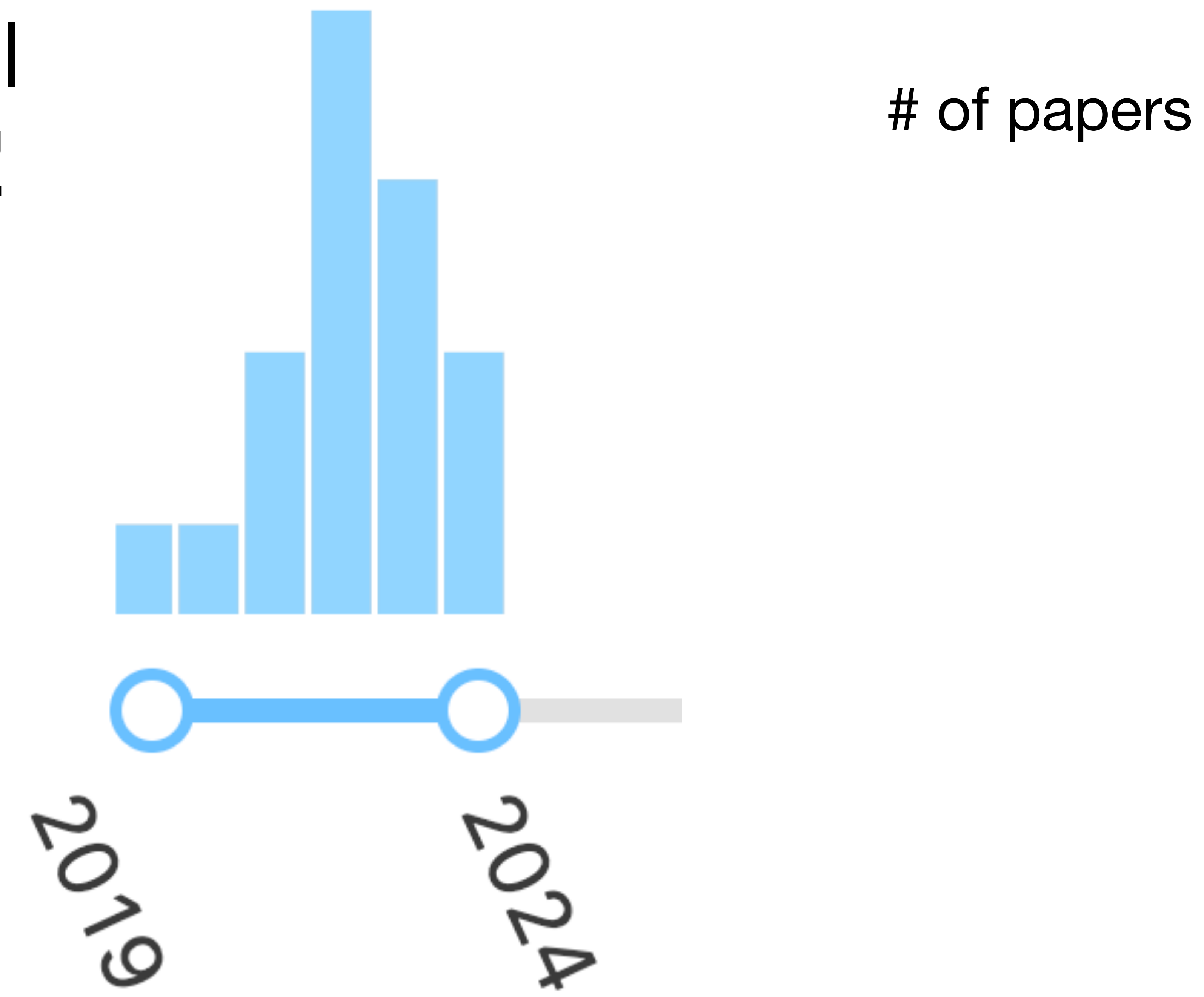
32 total results!



Inspire HEP search results for "machine learning RHIC"

&

20 total results!



Inspire HEP search results for "machine learning EIC"

ML is a rapidly growing field at RHIC and beyond!

Roadmap

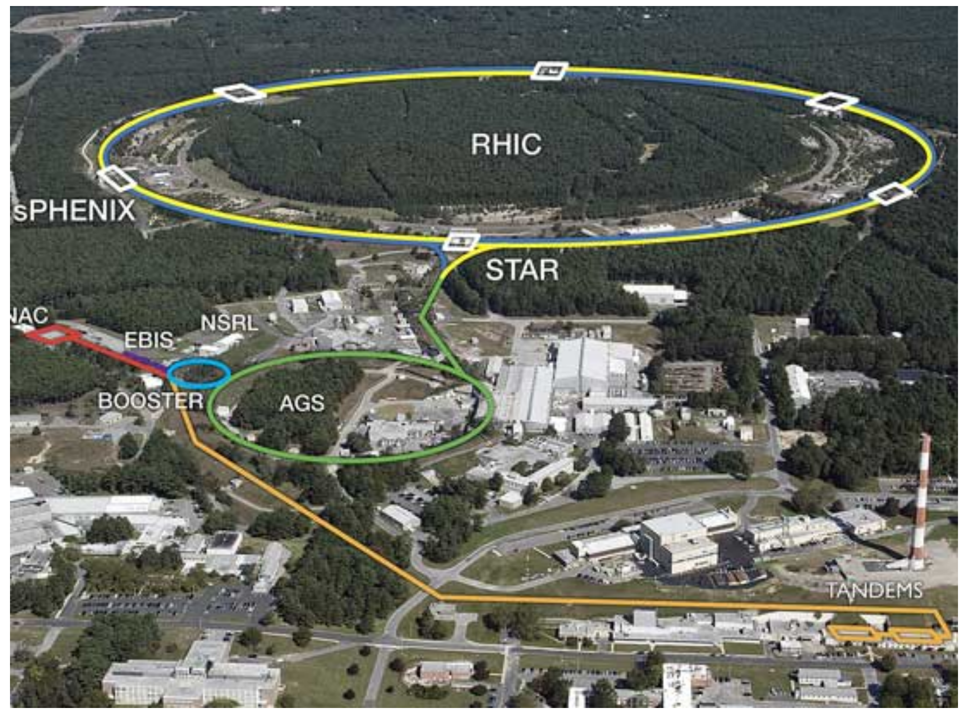
What is AI/ML and why is it useful for physics?

How is AI/ML currently being used at RHIC?

What are some uses of AI/ML beyond RHIC?

Data pipeline

ML is used throughout the data analysis pipeline!



RHIC Accelerator Complex → Experiments → Event Filtering → Reconstruction

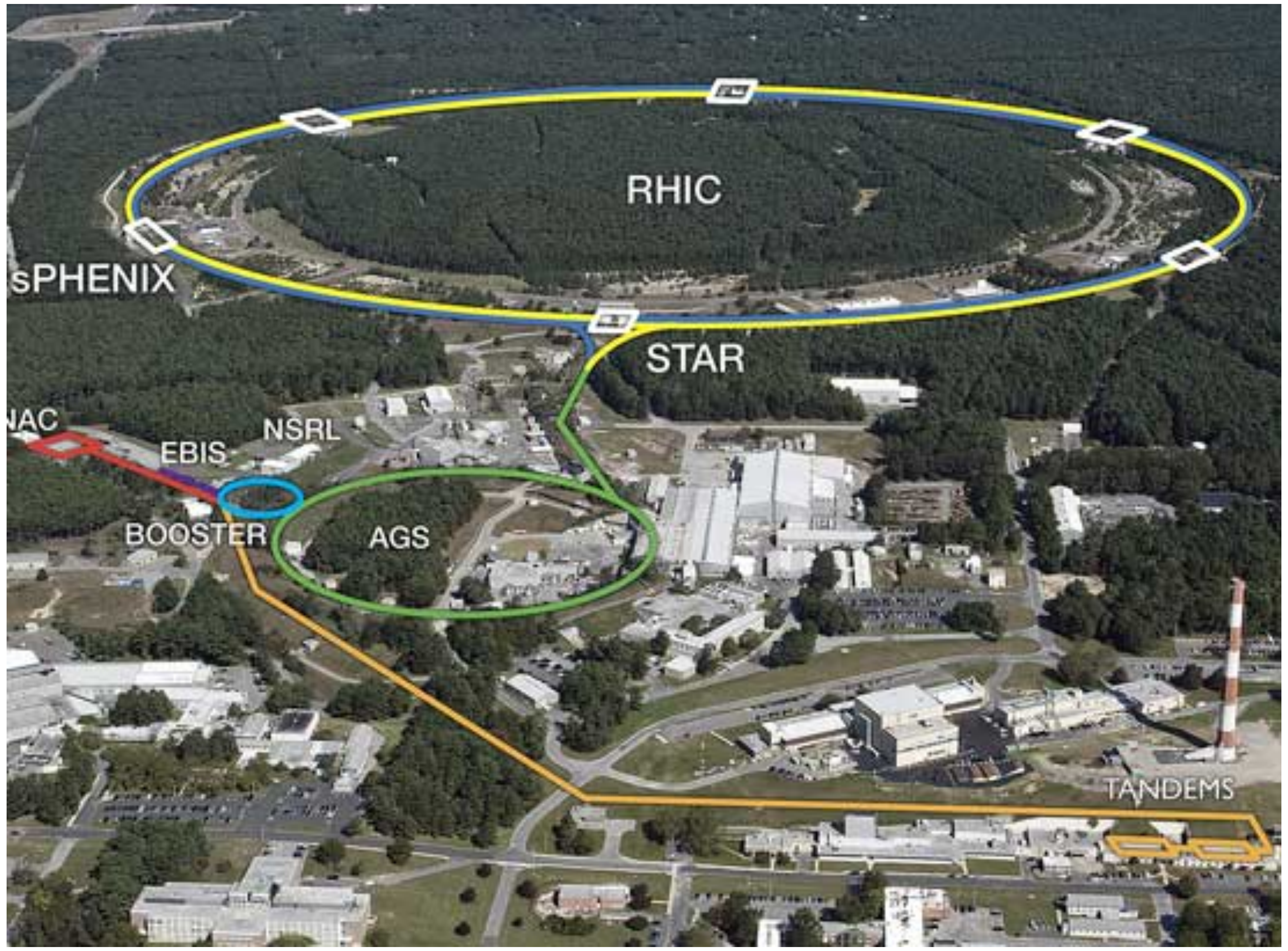
Reconstructed data

Simulation

Results ← Event/Analysis Selections ← Reconstructed data / Simulation

See Yeonju Go's talk today @ 9:30 am

ML @ the RHIC Accelerator Complex



- Boosted Decision Trees to identify and predict magnet quenches from historical data.
  - Combined with Autoencoders used to identify signs indicative of future quenches.

[JACoW IPAC2023 (2023) WEPA10]

- Autoencoders and PCA used for dimensionality reductions to see which parameters are useful for beam cooling. [JACoW NAPAC2022 (2022) 260-262]

- Algorithms need to be robust to machine parameters.
  - Reinforcement or unsupervised learning useful.
- Need machine development time, can use simulations.

[JACoW ICALEPCS2023 (2023) FR2AO04]

See Xiaofeng Gu's talk today @ 2:30pm and Yuan Gao's talk today @ 3:30 pm

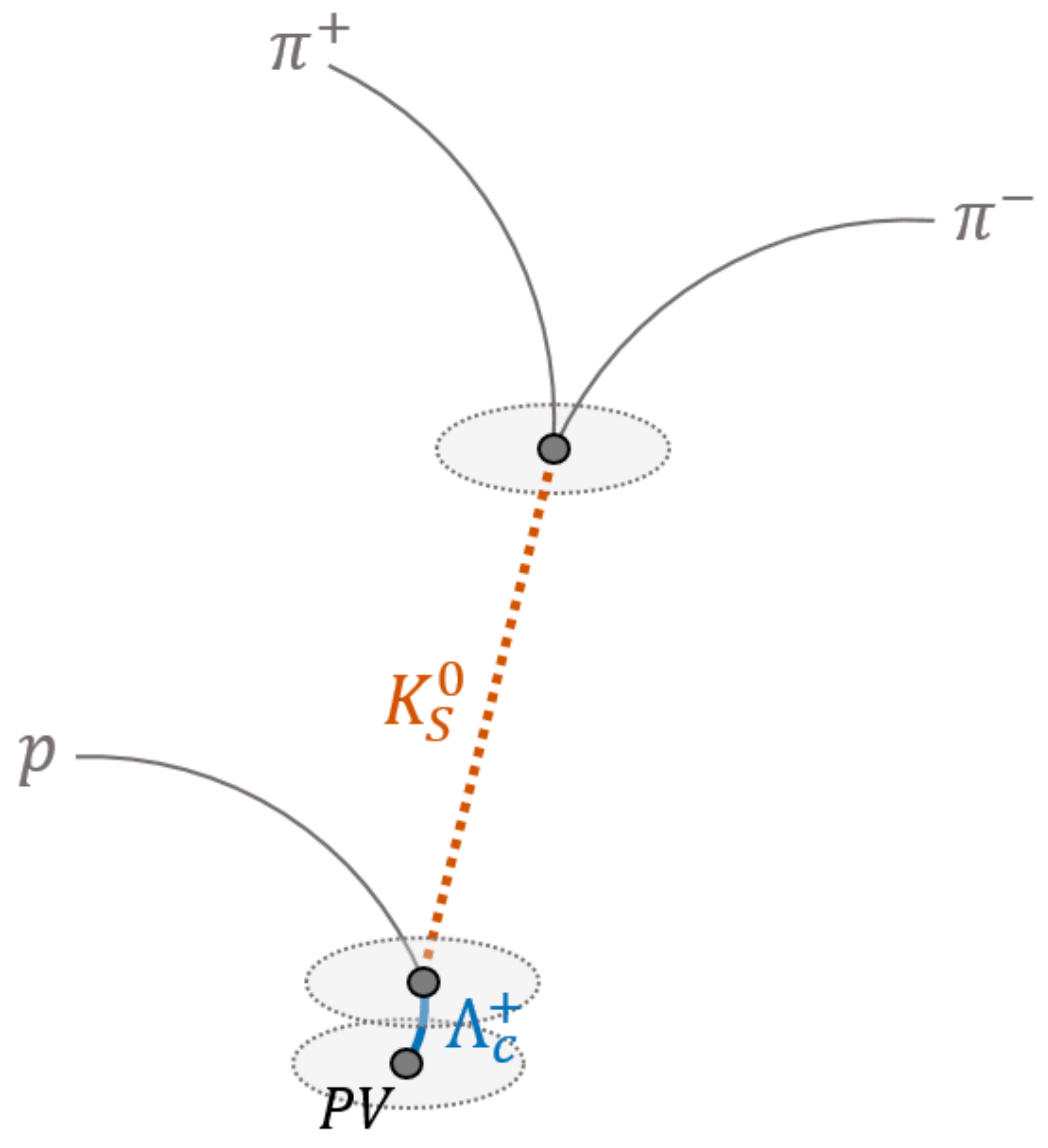
Event filtering

- Data volume is increasing at a fast rate, need solutions for limited computing resources.
  - Raw volume up to PB/s! If we took all raw data, would easily exceed storage capabilities.
- Two potential ways machine learning can help!
  - **Solution #1:** Perform fast selection/rejection of data with ML integrated into the firmware using FPGAs
    - Employs high level synthesis package hls4ml See Cameron Dean's talk today @11:30 am
  - **Solution #2:** Reduce data size
    - Very important for the case of storing all viable collisions (streaming readout)
    - Autoencoders are natural data reducers!
    - Application in sPHENIX TPC [arXiv:2310.15026]

      See Yi Huang's talk today @11:00 am

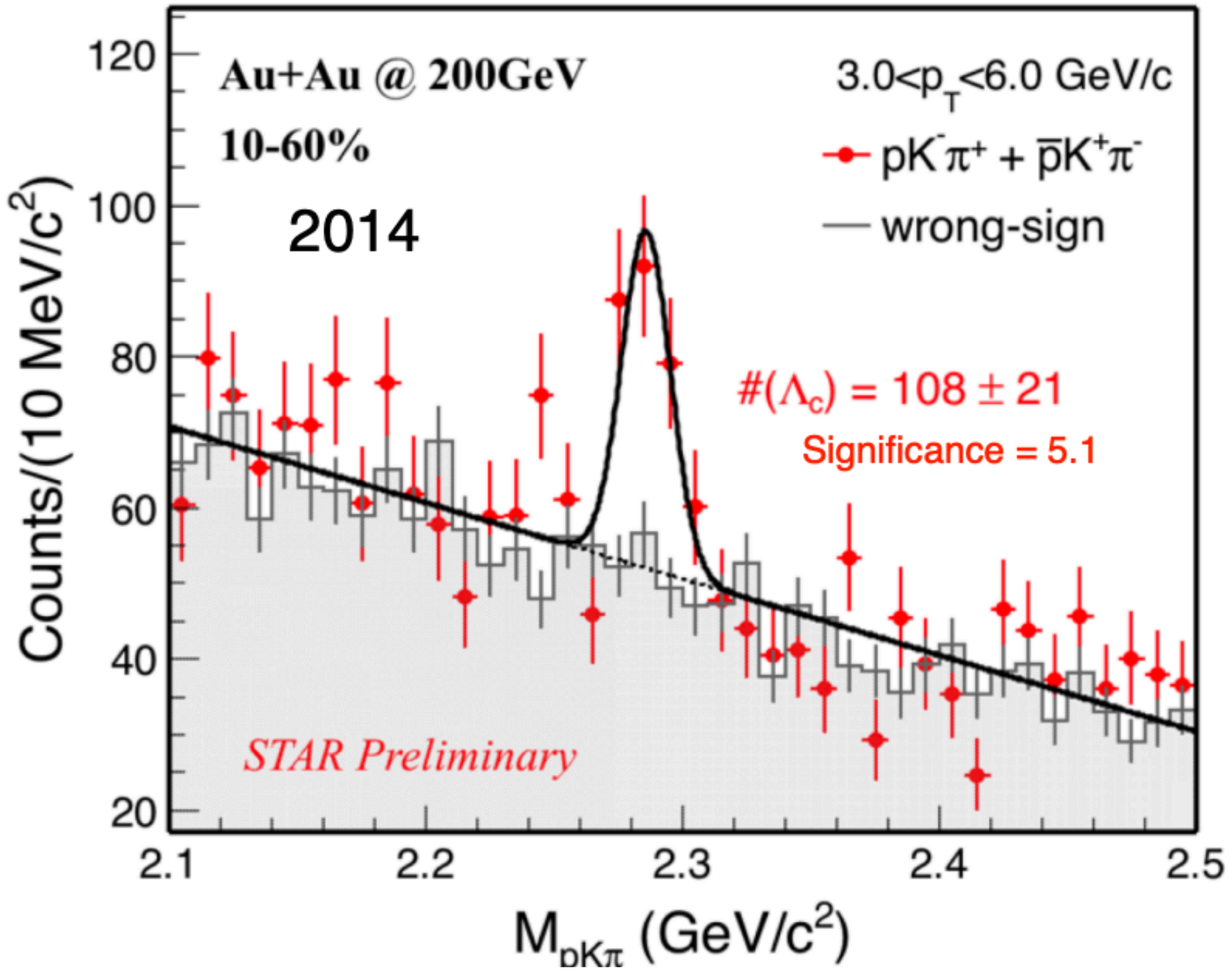
Signal/background discrimination

- ML has also seen a lot of success for applications in analysis!
- **Conventional approach:** Apply cuts to tag particle based on decay topology
  - Becomes difficult in heavy-ion environment with a large background.
- **ML-Based approach:** Use low-level parameters such as constituents, secondary vertices, track impact parameters etc. Learn from simulation in a supervised approach.
- Well established at RHIC and the LHC, many success stories!

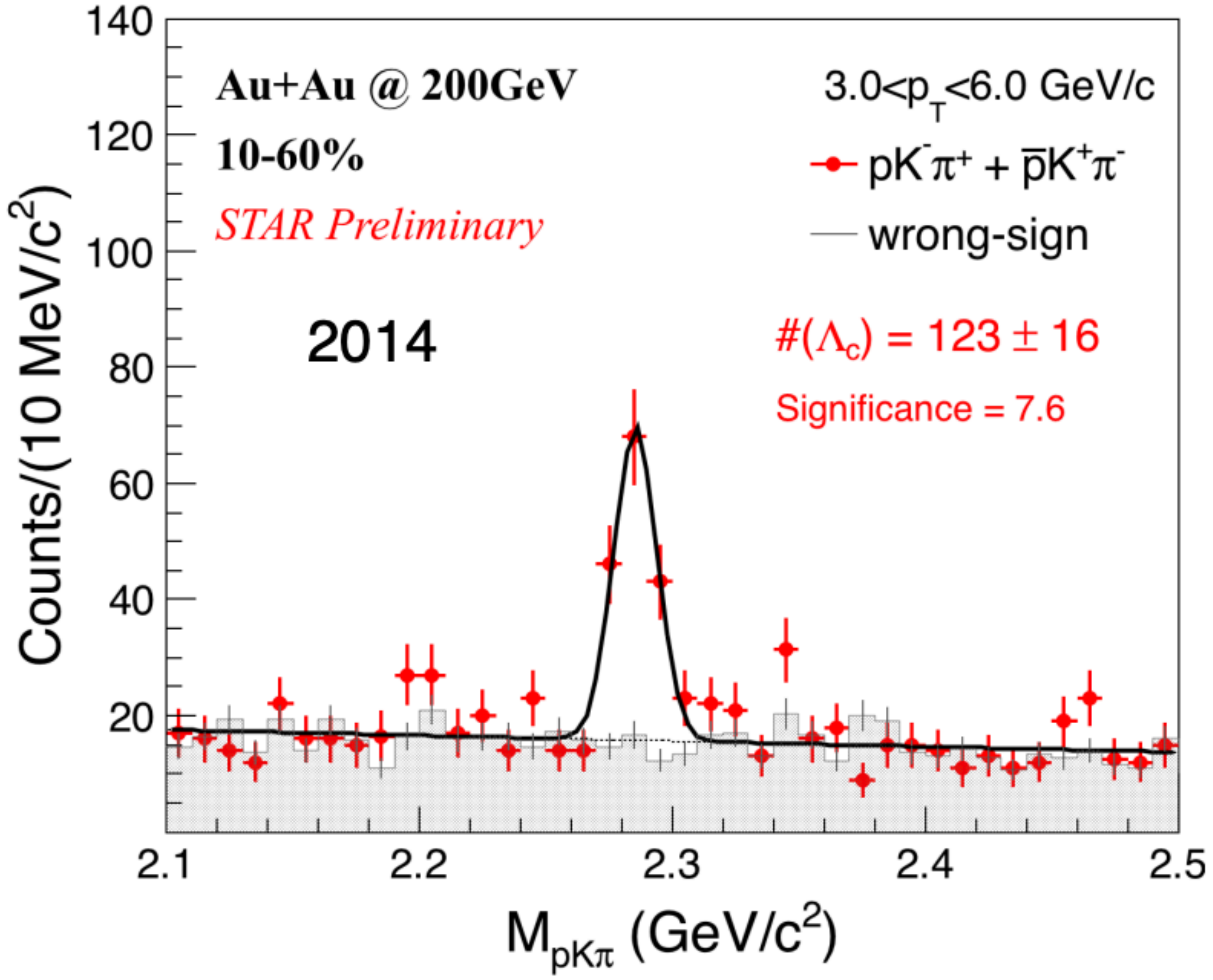


Signal/background discrimination

Traditional Techniques



With BDT

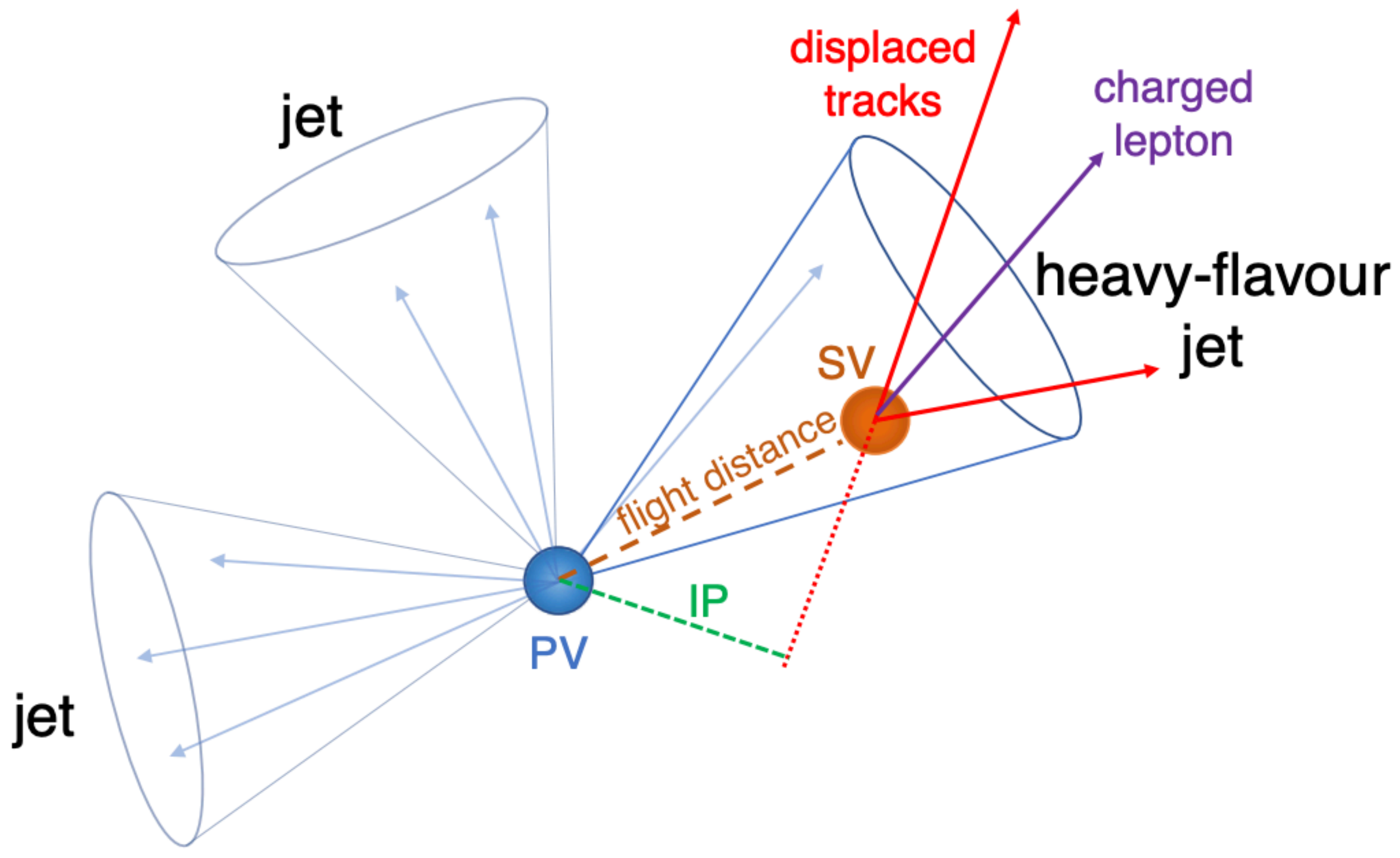


STAR

[PRL 124, 172301 (2020)]

- Boosted Decision Tree implemented in ROOT TMVA to optimize signal for Λ_c baryon production.
- Trained in a supervised manner with EvtGen
- 50% increase in signal significance with ML!

Heavy flavor jet tagging



JINST 13 (2018) 05, P05011

Goal: identify jets initiated by a heavy-quark (HF jet)

Conventional approach: Apply cuts to select jets with displaced decay vertices and large impact parameter tracks.

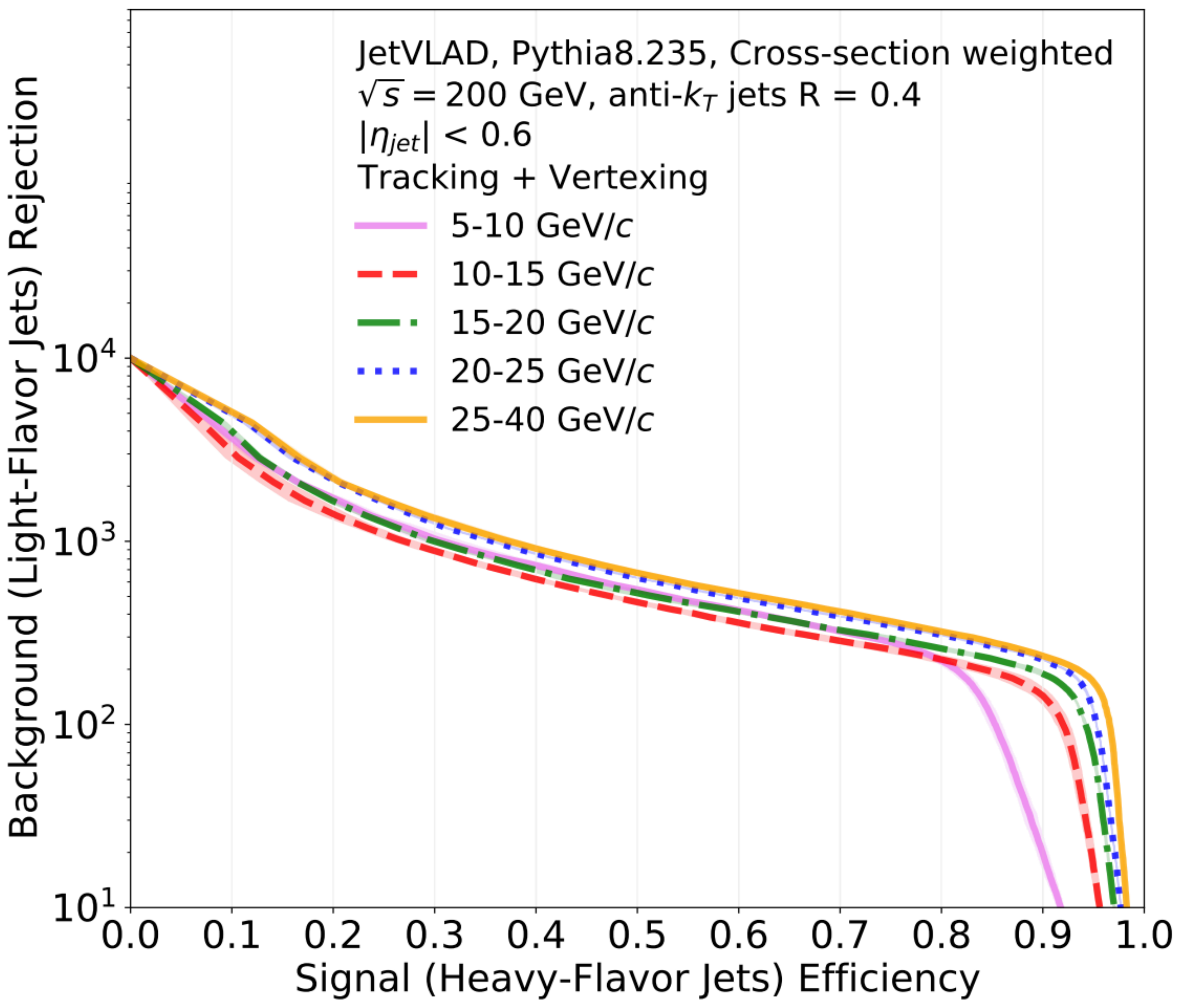
ML approach: Use low-level jet parameters such as constituents, secondary vertices, track impact parameters etc to learn from simulation in a supervised approach.

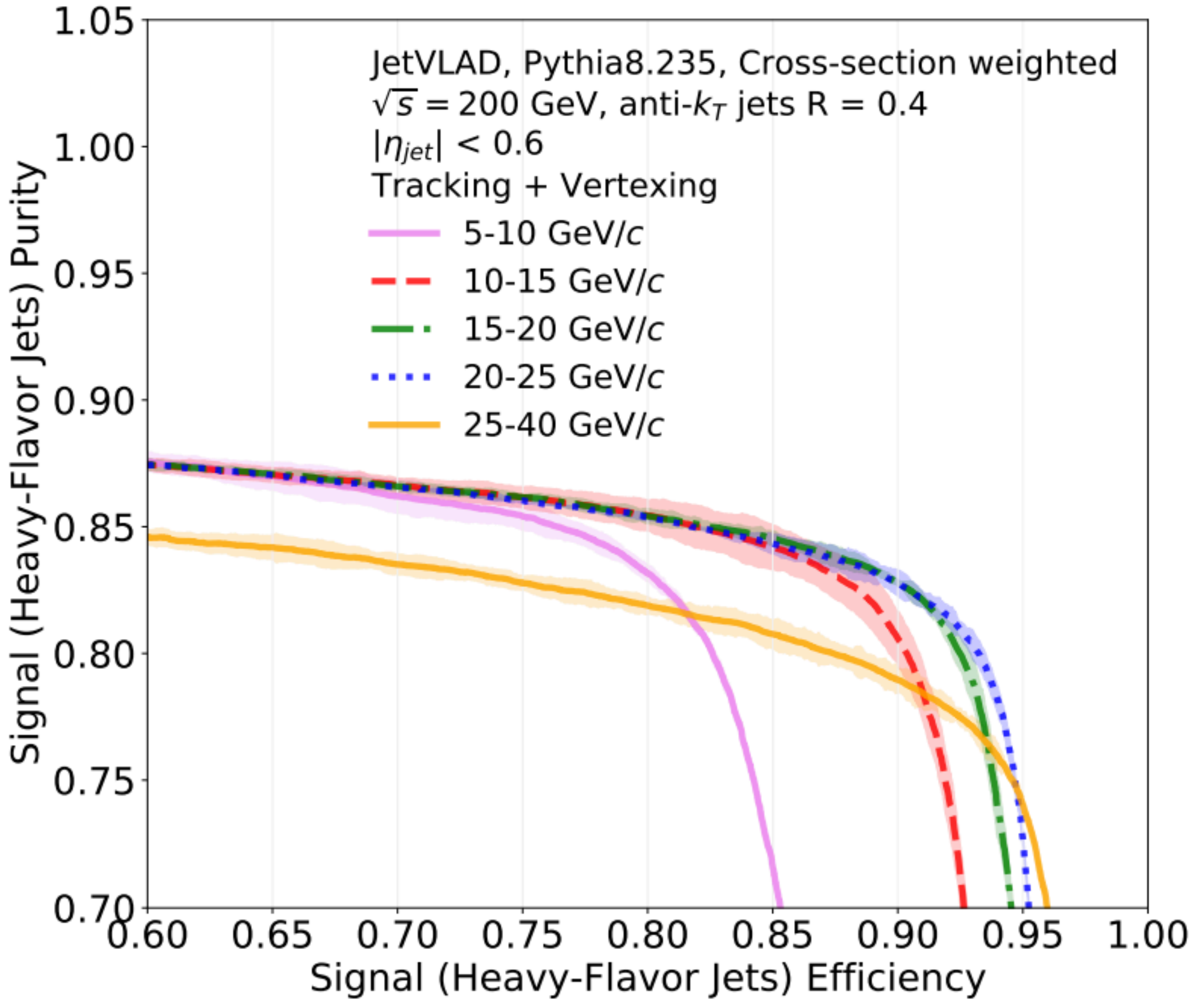
Look at dependence of parton mass on parton shower + hadronization in vacuum and in medium!

Jet VLAD [JINST 16 (2021) 03, P03017]

For input to the model treat the jet as a set of particles $\mathcal{J} = \{(p_{T,i}, \eta_i, \phi_i, \ldots)\}_{i=1}^{n}$

Model includes pooling layer that takes set of feature descriptors as an input and returns a fixed-length feature vector that characterizes each set.



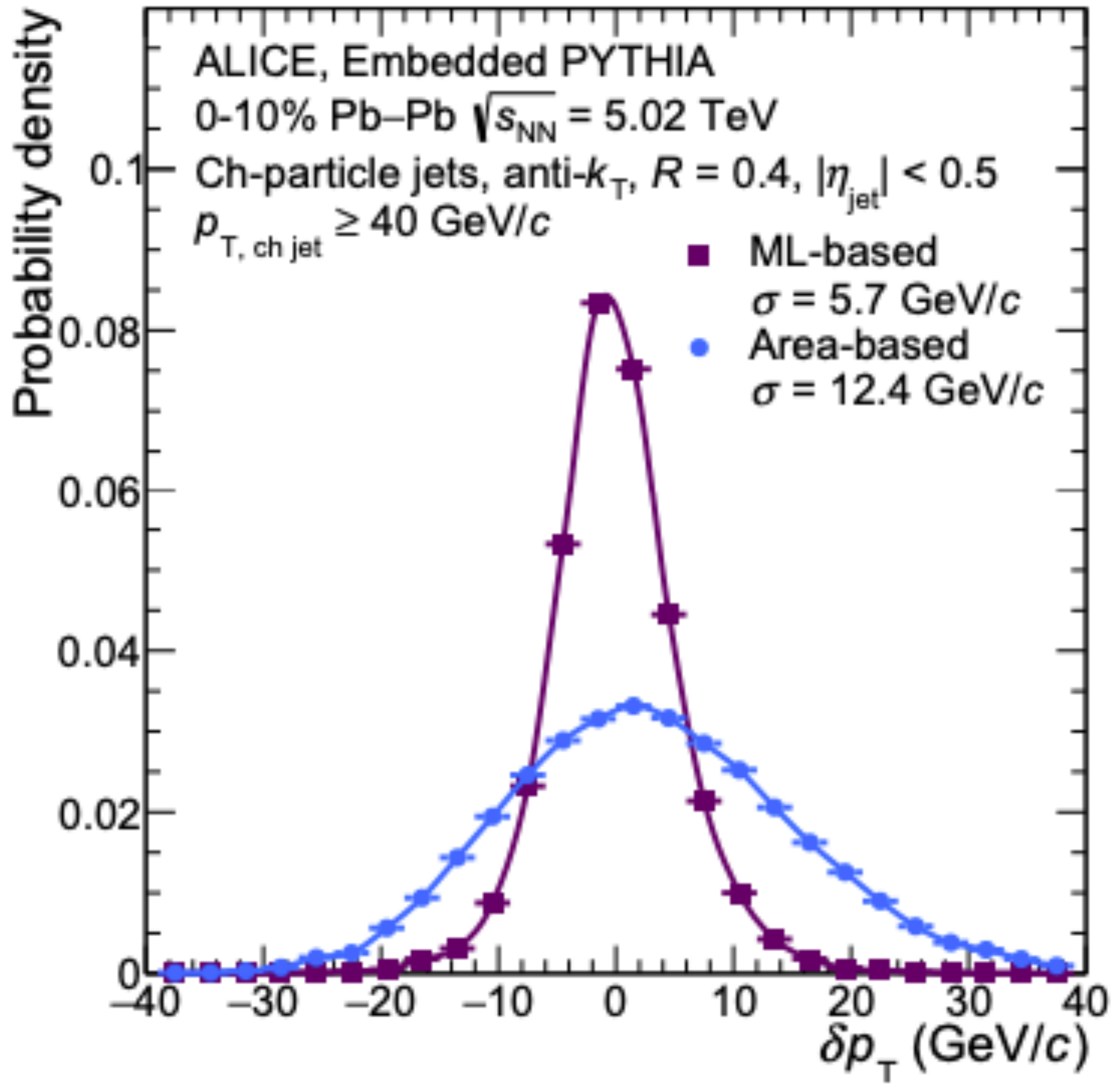


- For higher p_{T} HF jets, background rejection increases, but purity decreases
- Fragmentation changes as function of p_{T} leads to an overlap of feature space

This is a challenging problem! Especially in Au+Au!

Jet background correction

[PLB 849 (2024) 138412]



$$\delta p_T = p_{T,rec} - p_{T,true}$$

Goal: Use properties of the jet and its constituent to determine the background-corrected jet p_T. A few approaches for this!

Shallow Neural Network in scikit-learn (simple tools) trained on PYTHIA embedded into HI background [PRC 99, 064904 (2019)]

Use interpretable ML to create methods to distinguish signal from the background [arXiv:2402.10945]

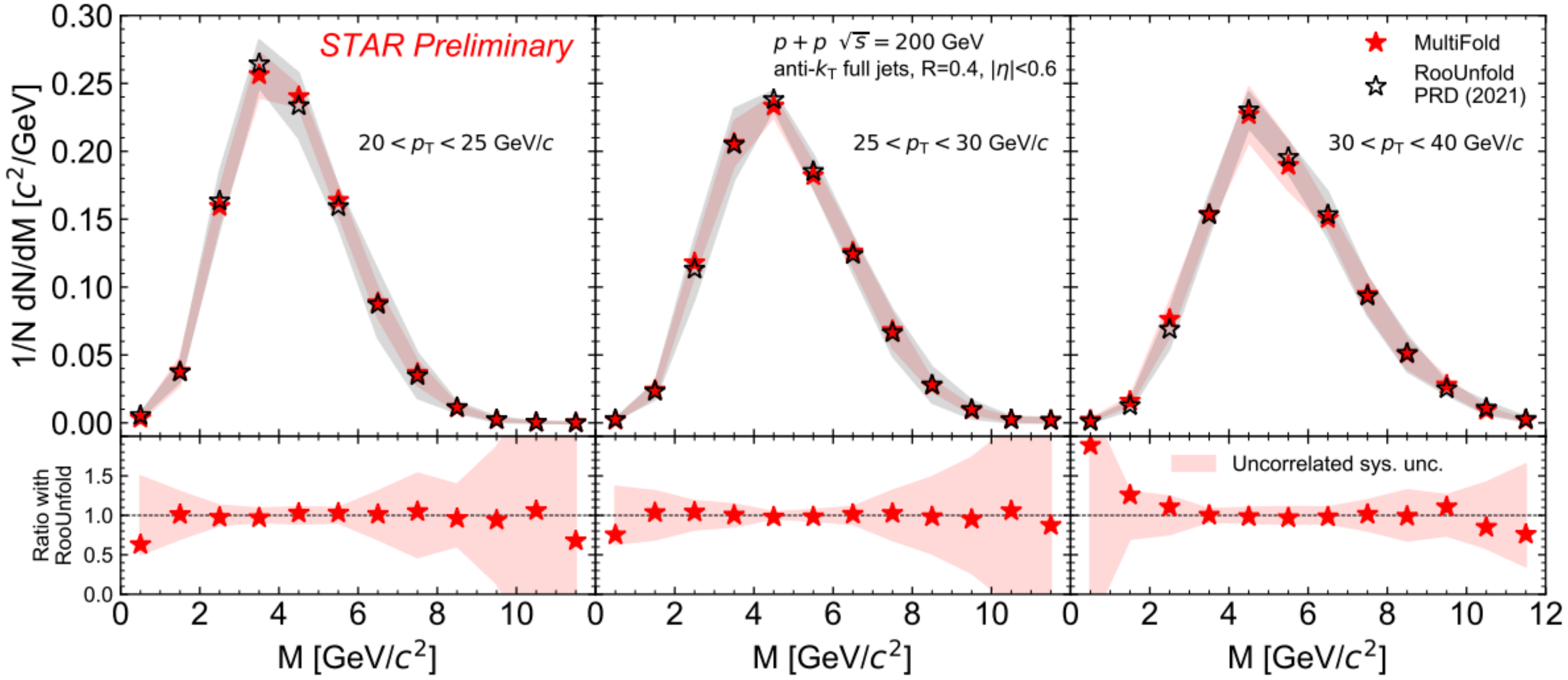
See Charles Hughes's talk today @ 10:00 am and Yilun Wu's talk today @ 12:00pm

Ongoing work to apply such techniques at RHIC! Stay tuned!

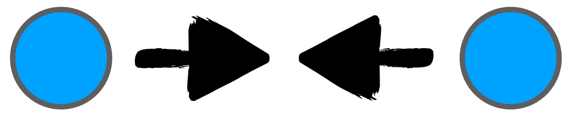
Unfolding with ML

- **Conventional Approach:** Apply unfolding procedure on a binned distribution and repeat for each observable.
- **ML-based Approach:** Use ML to calculate weighting factors and unfold the phase space all at once, before the choice of binning or observable! [PRL 124, 182001 (2020)]

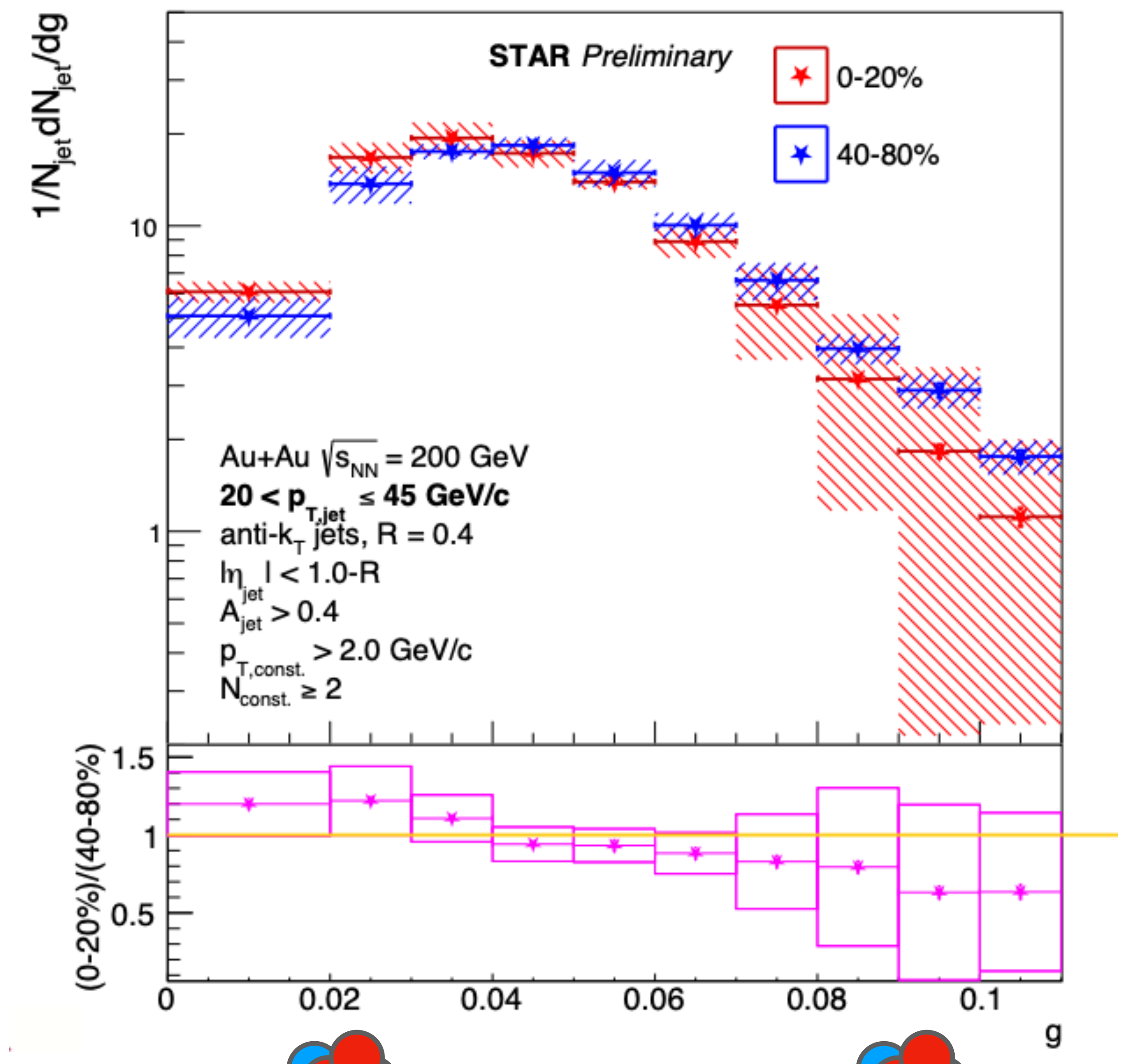
Has been applied in pp and Au+Au* at RHIC!

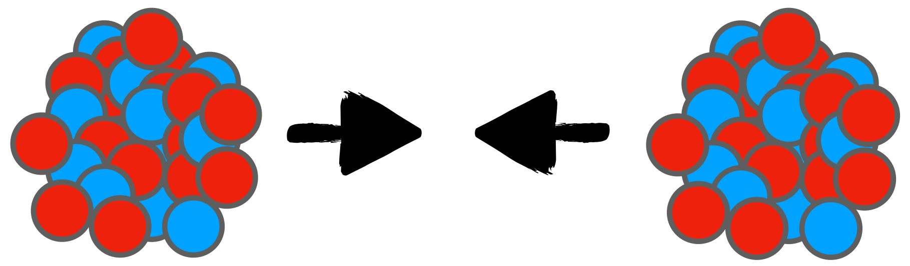


Youqi Song, DIS 2023, [arXiv:2307.07718]



Tanmay Pani, QM 2023 [arXiv:2403.13921]





See Youqi Song's talk today @ 4:00 pm and Hannah Harrison's talk today @ 4:30

** model uncertainty not yet evaluated in Au+Au*

Roadmap

What is AI/ML and why is it useful for physics?

How is AI/ML currently being used at RHIC?

What are some uses of AI/ML beyond RHIC?

Where are we going?

Very large volumes of will be taken and analyzed in the decades to come - new tools will be *increasingly important*!

ML at the EIC

Electron Ion Collider is a future facility being designed with future techniques in mind!

Ongoing Activities w/ AI

- Detector design
- Simulation
- Reconstruction See Derek Anderson's talk @ 2:00 pm
- Particle Identification
- Analysis

See [AI4EIC] for a comprehensive overview

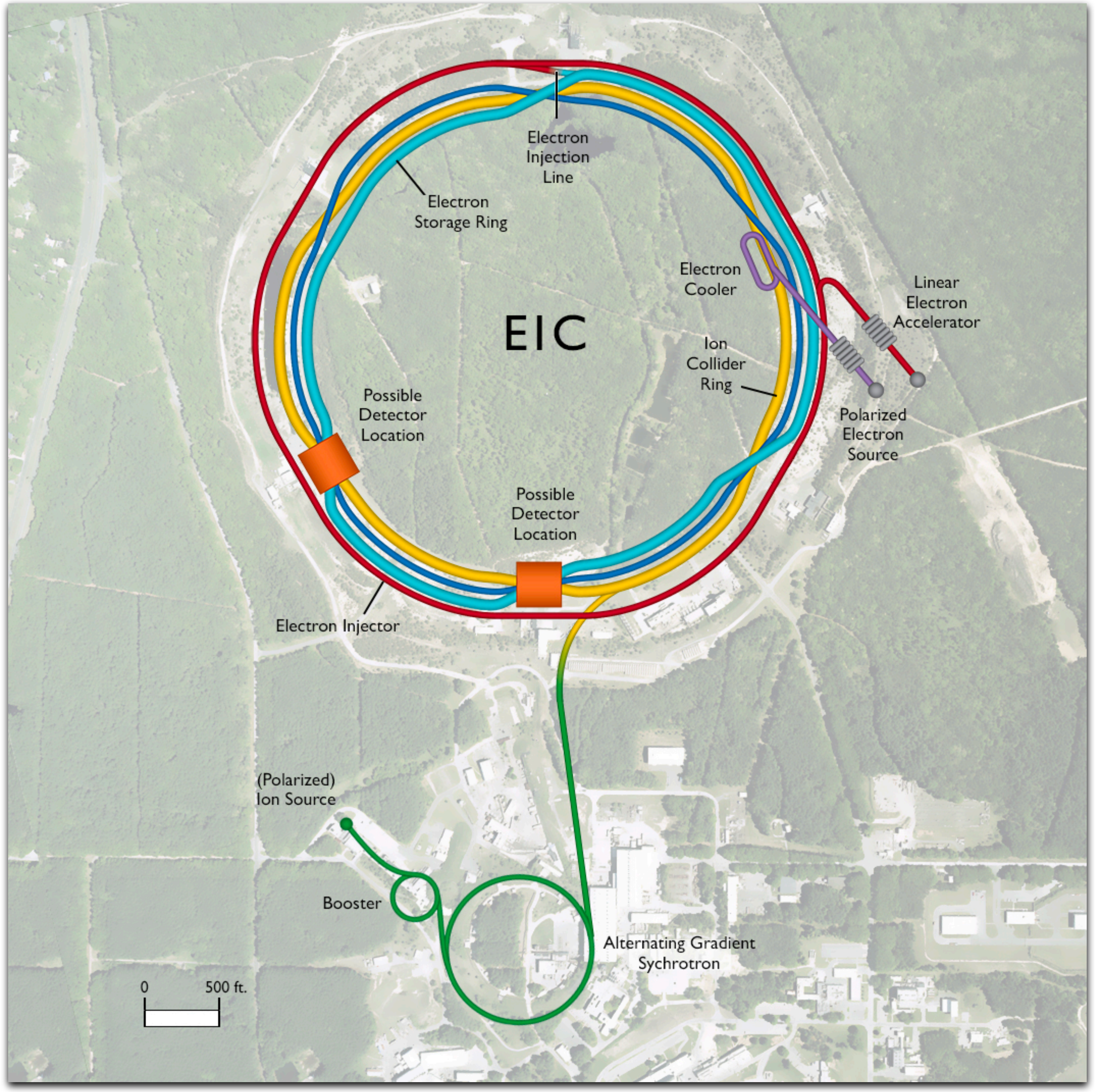


Image Credit: [Brookhaven National Lab]

See Cristiano Fanelli's talk @ 1:30pm

Event classification [JHEP 03 (2023) 085]

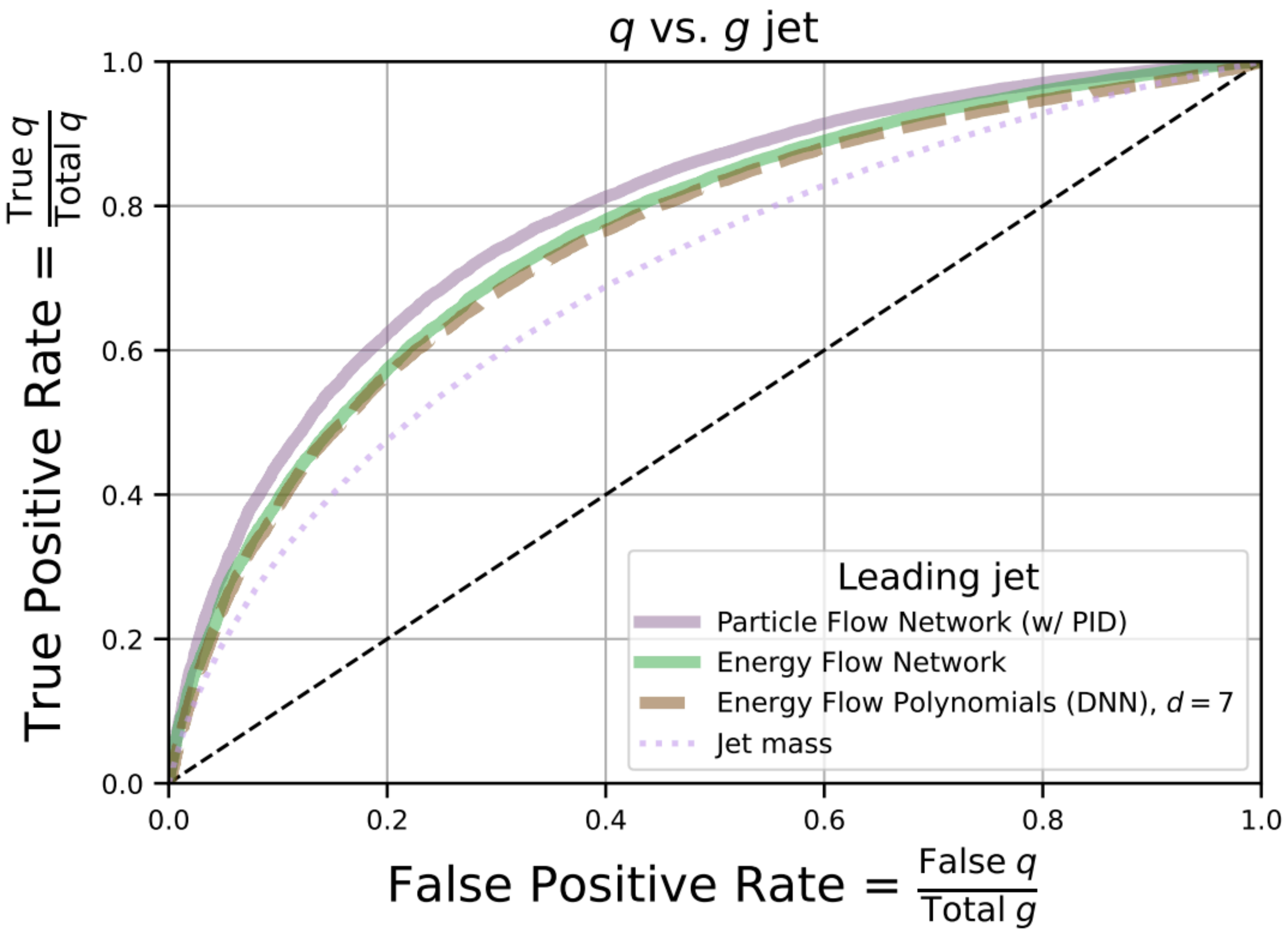
- Study the effectiveness of ML-based classifiers to
  - Identify the flavor of the jet
  - Identify the underlying hard process of the

- Additionally study the effectiveness of different ways of representing information
  - Particle Flow Networks [JHEP 01 (2019) 121]

$$f(p_1, \dots, p_N) = F\left(\sum_{i=1}^{N} \Phi(p_i)\right) \qquad p_i = (z_i, \eta_i, \phi_i, \mathrm{PID}_i)$$

  - Energy Flow Polynomials [JHEP 04 (2018) 013]

$$\mathrm{EFP}_G = \sum_{i_1} \cdots \sum_{i_V} z_{i_1} \cdots z_{i_V} \prod_{(k,l)\in E} \theta_{i_k i_l}$$



Indications that ML-based methods will have an improved performance over traditional techniques!

See also, [arXiv:2404.05752]

Open questions for next ~5 years

How do we assign a systematic uncertainty for the ML?

How do we construct more interpretable models?

How can we ML-based applications reproducible?

Do we need to standardize ML applications across experiments?

What is beyond this time scale?

ML for underlying physics

Could we use ML to directly access underlying physics mechanisms?

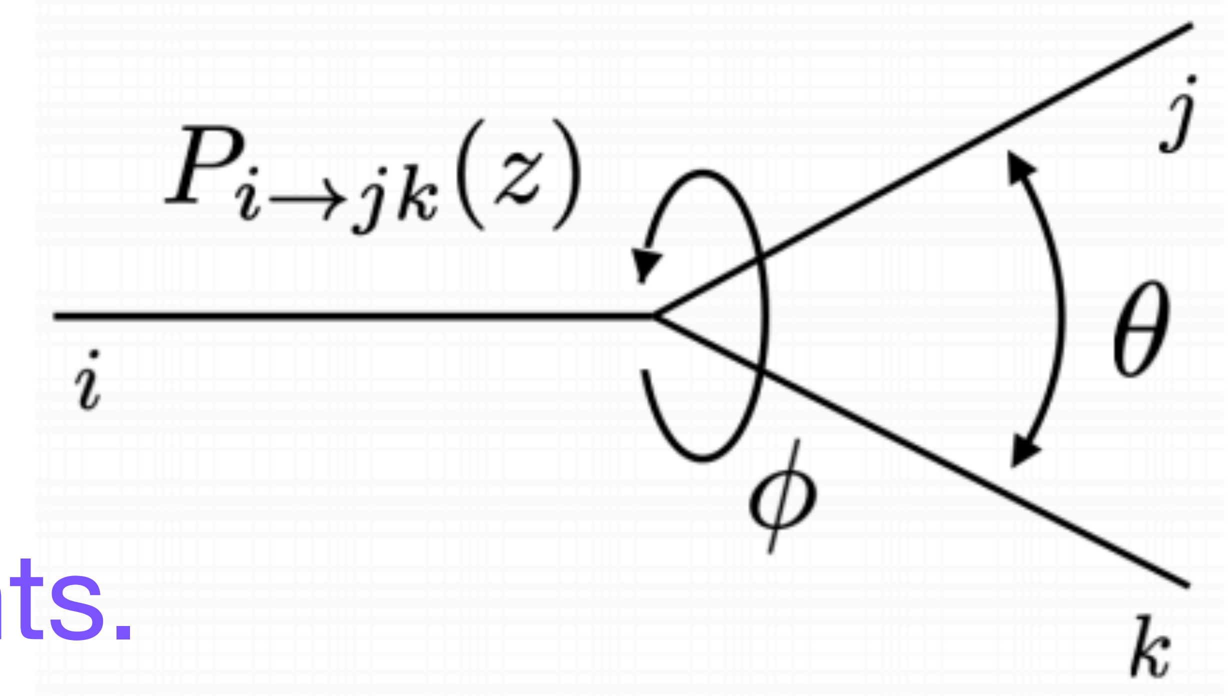
"Data"-based learning complements simulation-based inference.

~ Given an answer ~ "White Box" ML ~ Underlying physics	~ Domain knowledge ~ "Black Box" ML ~ Answer

This is a long term effort!

- Learning from data is difficult due to systematic experimental biases.
- Helpful in understanding uncertainties or shortcomings of models!

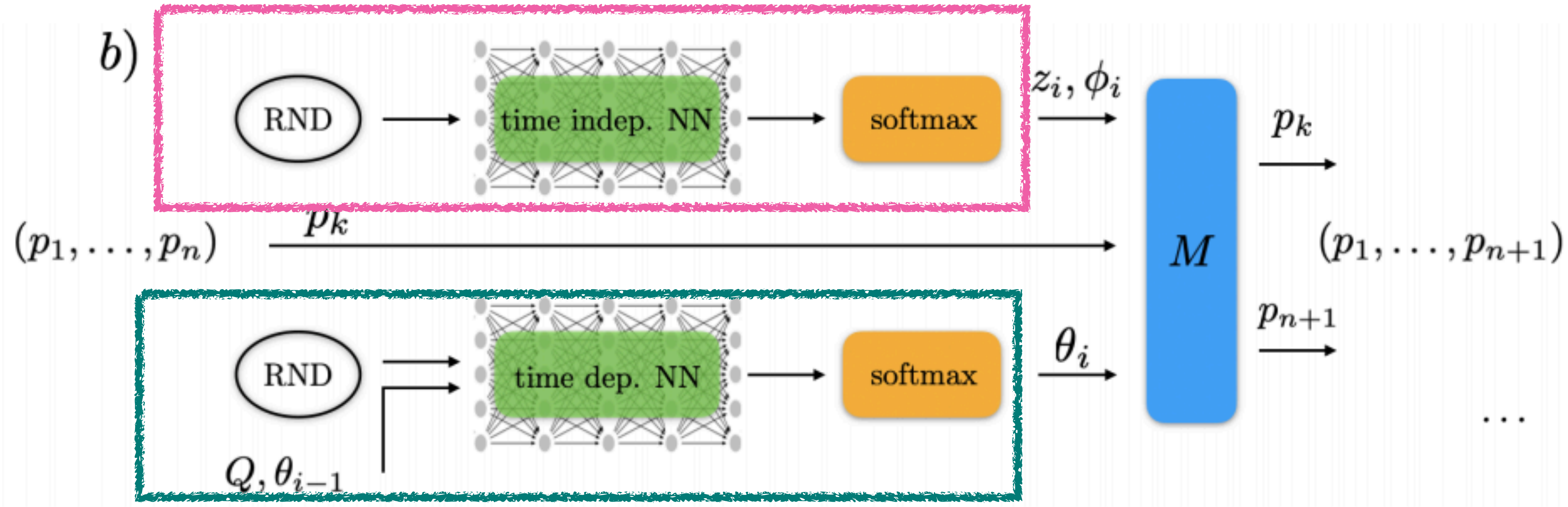
Proof of concept



- Extract splitting function from the network in white-box ML.

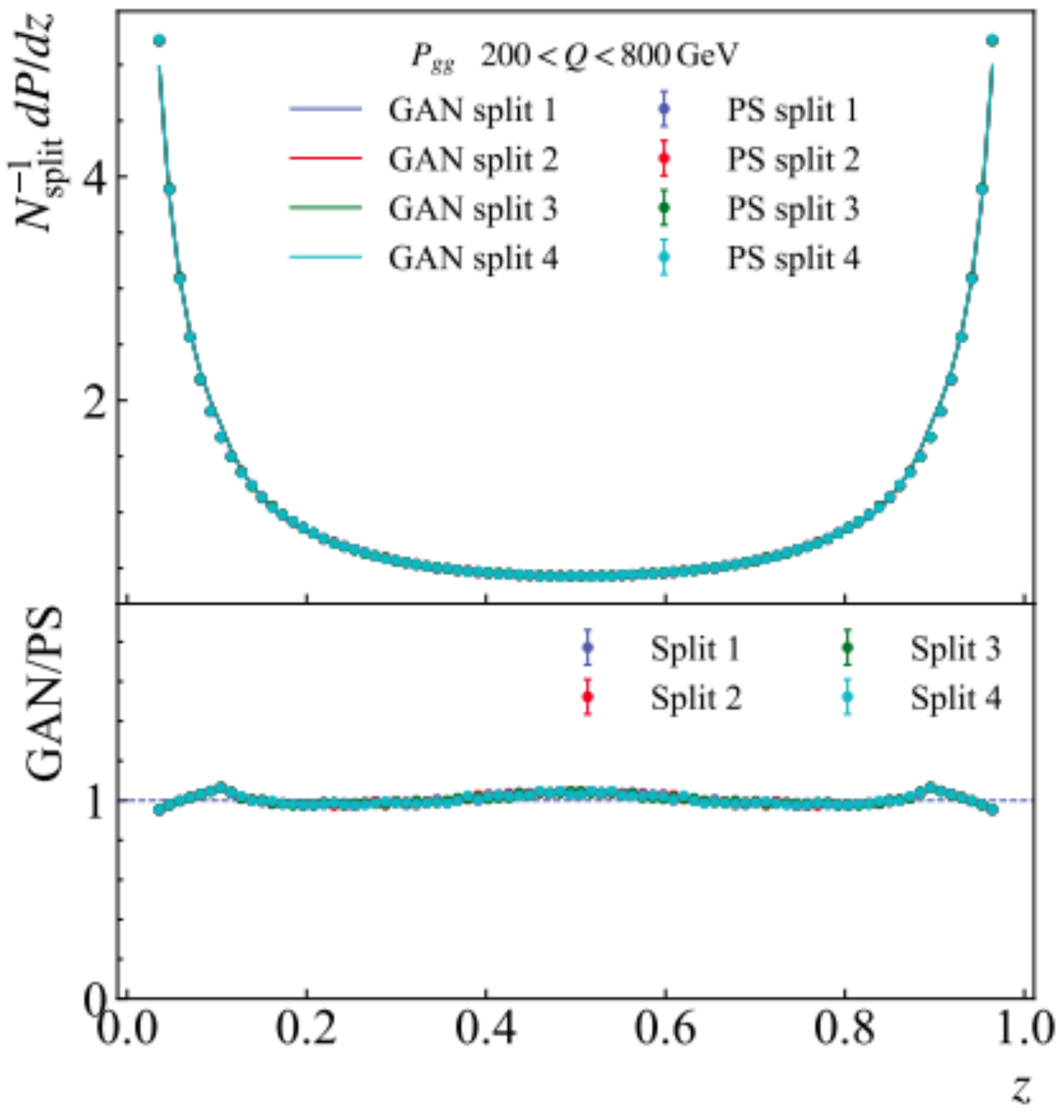
Done with a Generative Adversarial Network split into two components.

1. Time independent learns the z, ϕ



2. Time dependent learns the θ

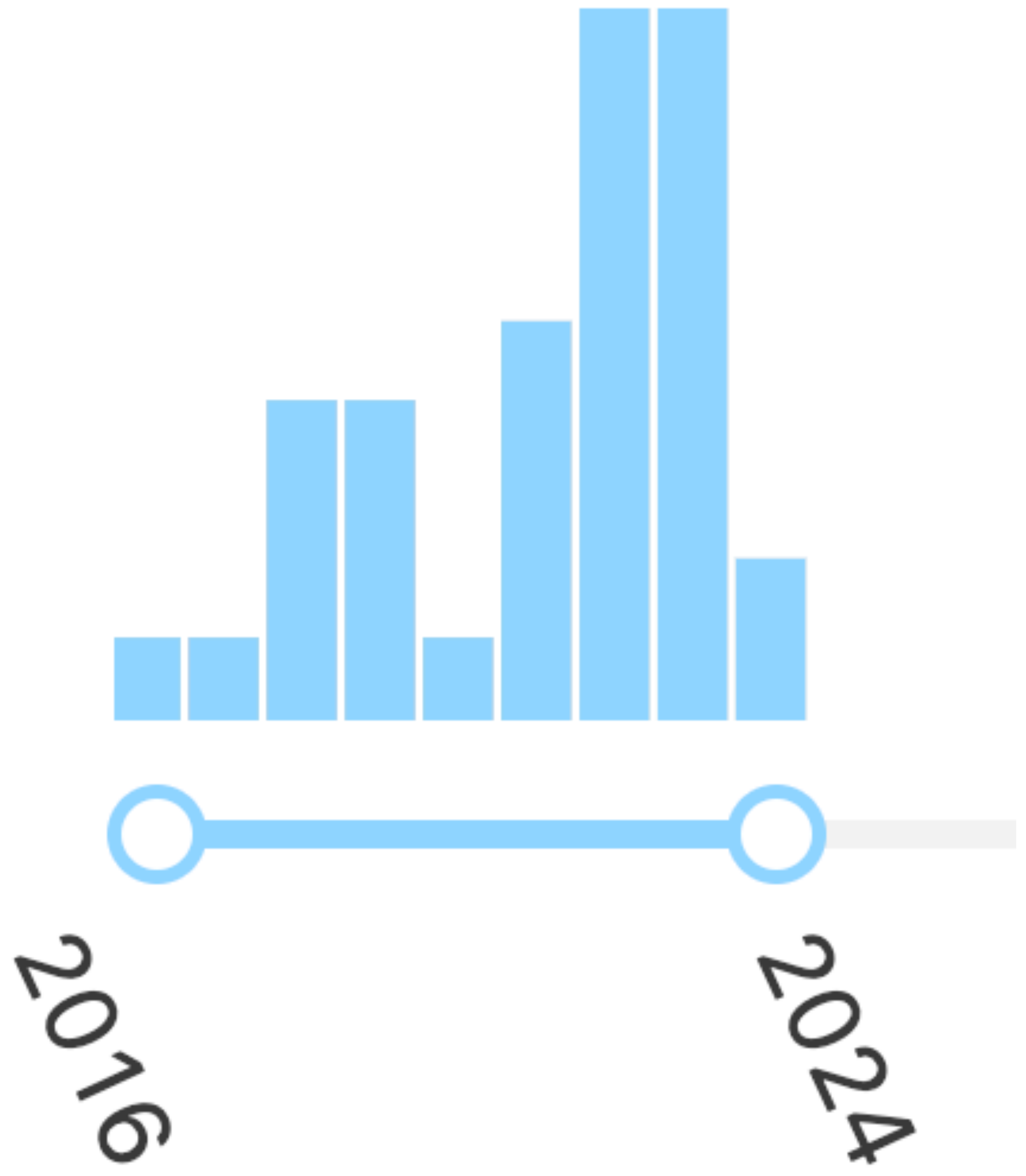
[PLB 829 (2022) 137055]



Was able to reproduce AP splitting function.

Conclusions

- We are taking more data with more complex measurements than ever before
- Machine learning and its use at RHIC is becoming increasingly more important
  - Will be crucial at the EIC!
- Lots of great experimental progress throughout the whole data analysis pipeline!
  - Talks today will cover this progress in great detail!



Future is very bright!

deep ai image editor

Backup

How does ML learn?

Supervised Learning

Algorithm learns from a labeled set of "true values".

Driven by the Task

Analogy: Taking a test

Unsupervised Learning

Algorithm finds structure in the data without knowing the desired outcome.

Driven by the Data

Analogy: Discovering allergies

Reinforcement Learning

Algorithm learns in a reward based system to determine a series of actions.

Driven by the Reward

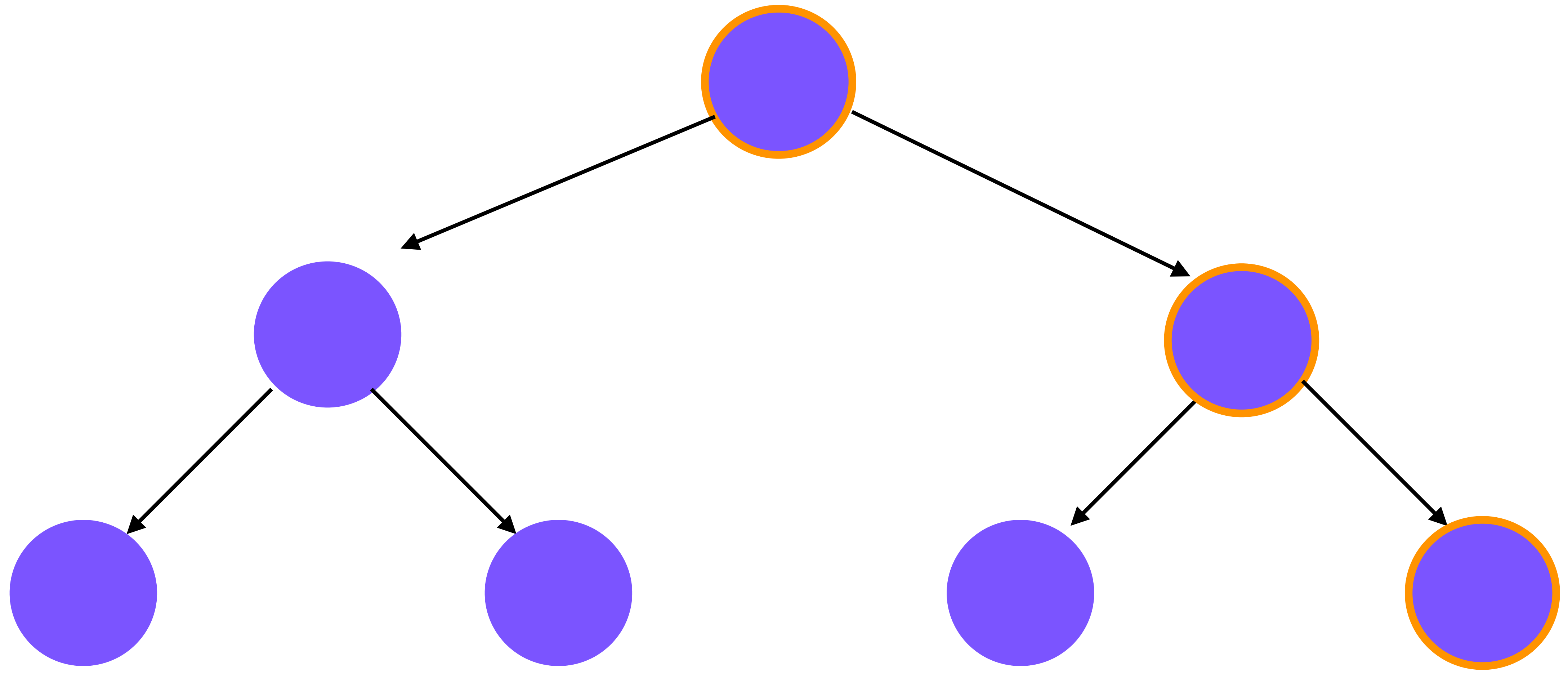
Analogy: Dog training

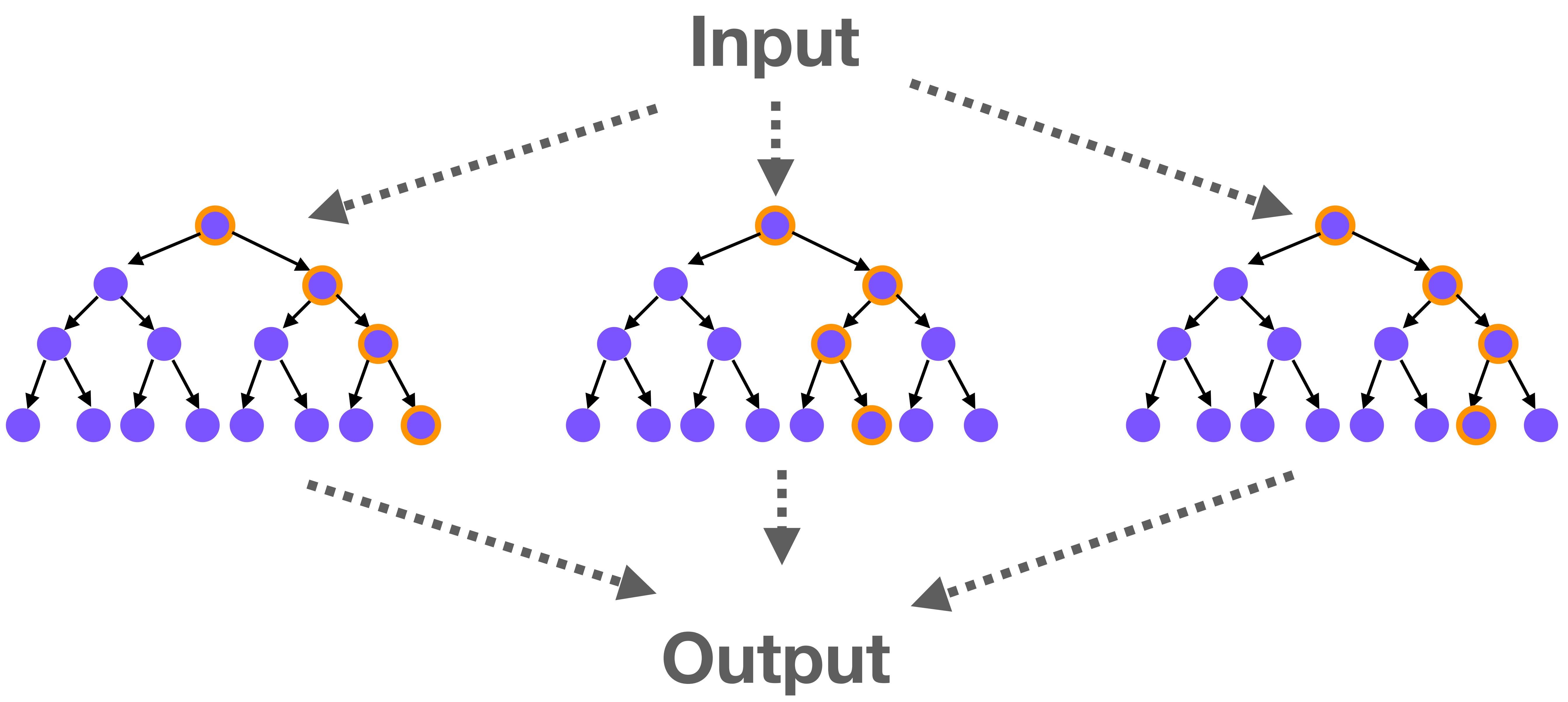
Intro to Random Forest

Random forests are composed of decision trees.

Decision trees are a set of rules organized in a tree structure.

Each node is a rule which subdivides the dataset into two or more parts (think 20 questions).





Output of the random forest is a combination of the output of each of the decision trees.

In training, the algorithm sets up the rules of each decision tree.

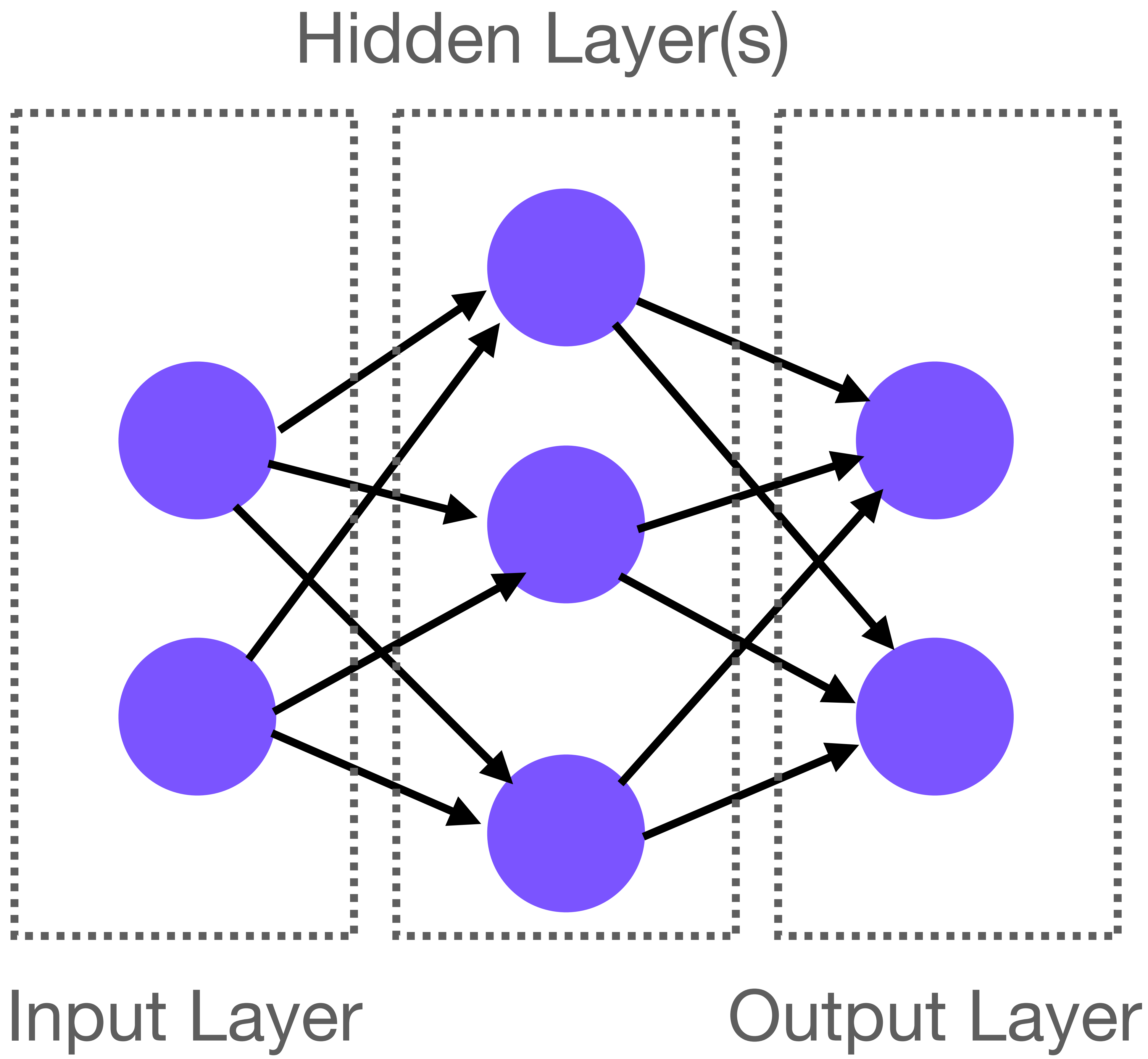
Neural Networks

Flow of information happens between nodes.

A weight is associated with each input to a given node.

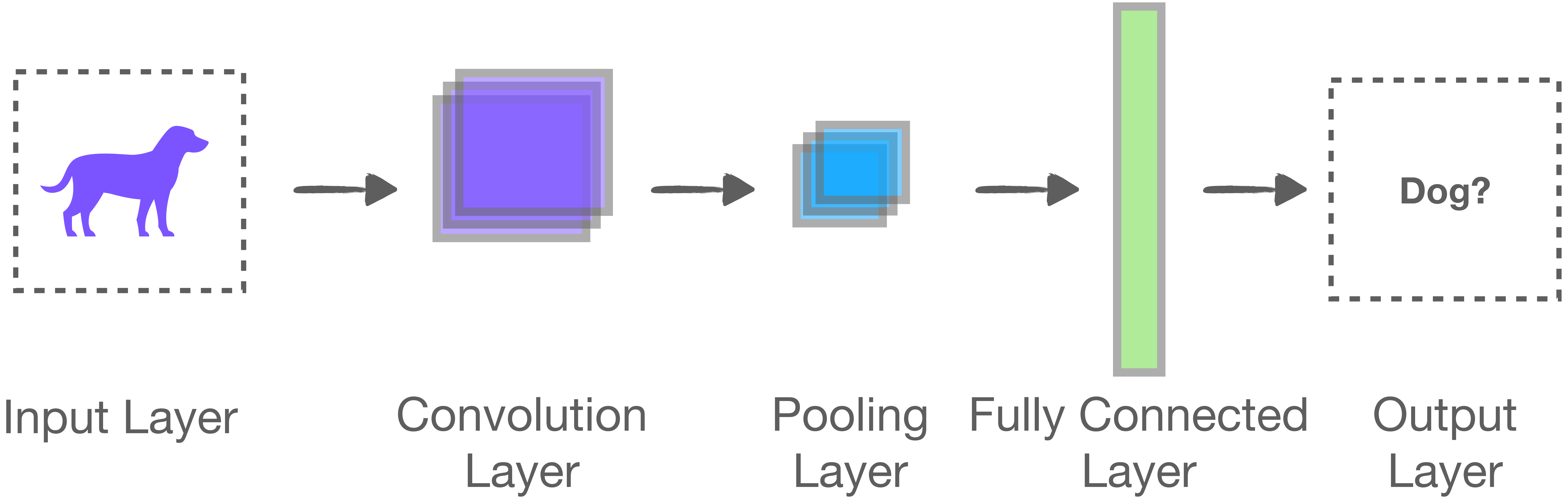
The output of each node is a function of the weighted inputs. The output of a node j, is generally written something like

$$O_j = \sum_{i=0}^{N-1} w_{ij} O_i$$

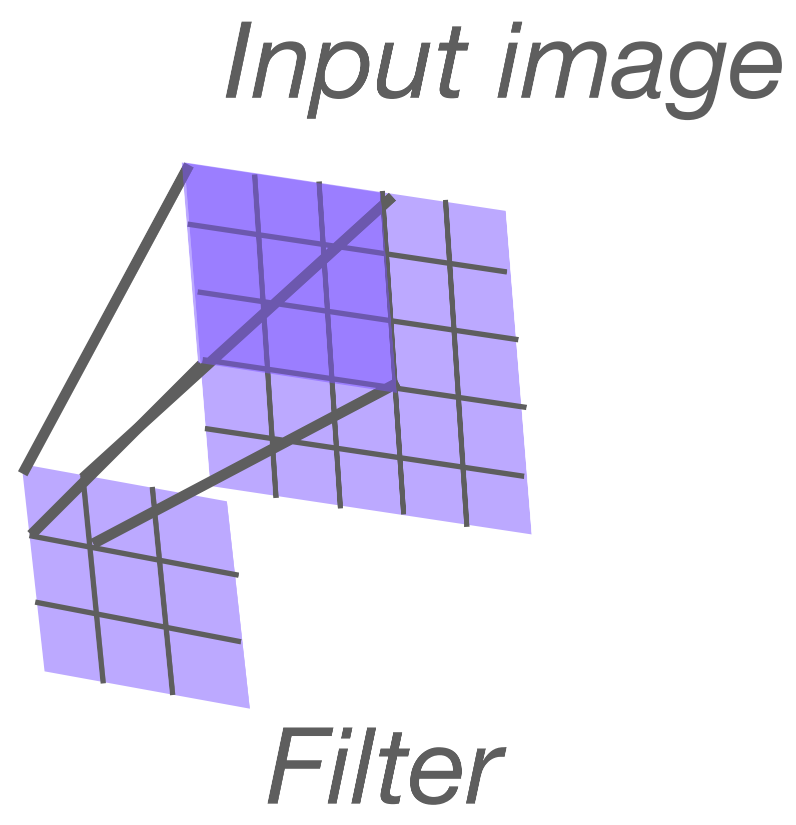


In training we seek to learn the set of weights which minimize the total error of the network.

Convolutional Neural Networks (CNNs)



→ Key component of a CNN is the **convolution layer**, which (with the help of a filter) will determine if a feature/pattern is present.



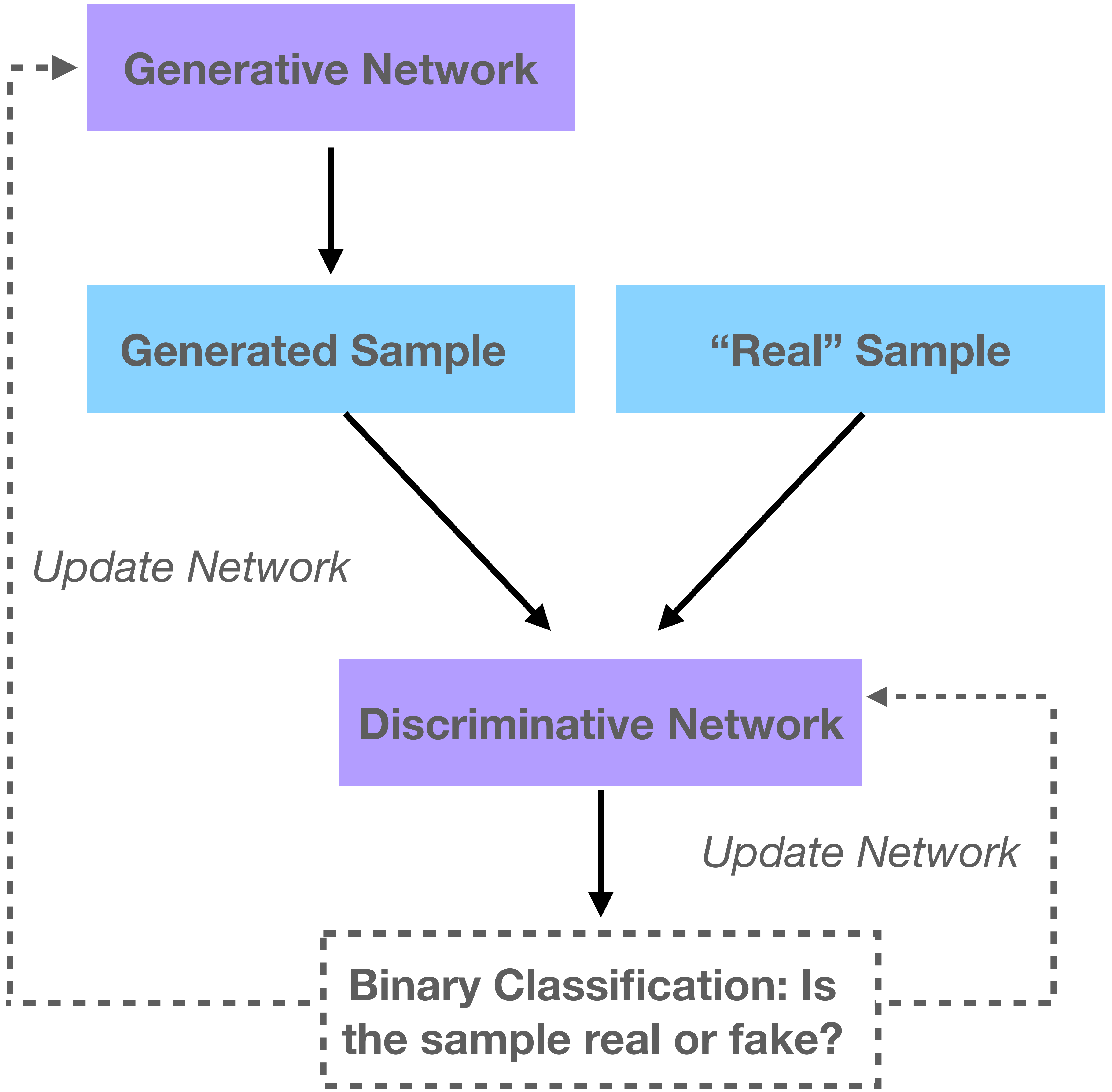
Generative Adversarial Networks (GANs)

Two networks compete with one another in a game.

The generative network seeks to fool the discriminative network.

The discriminative network seeks to find the real sample from the generated samples.

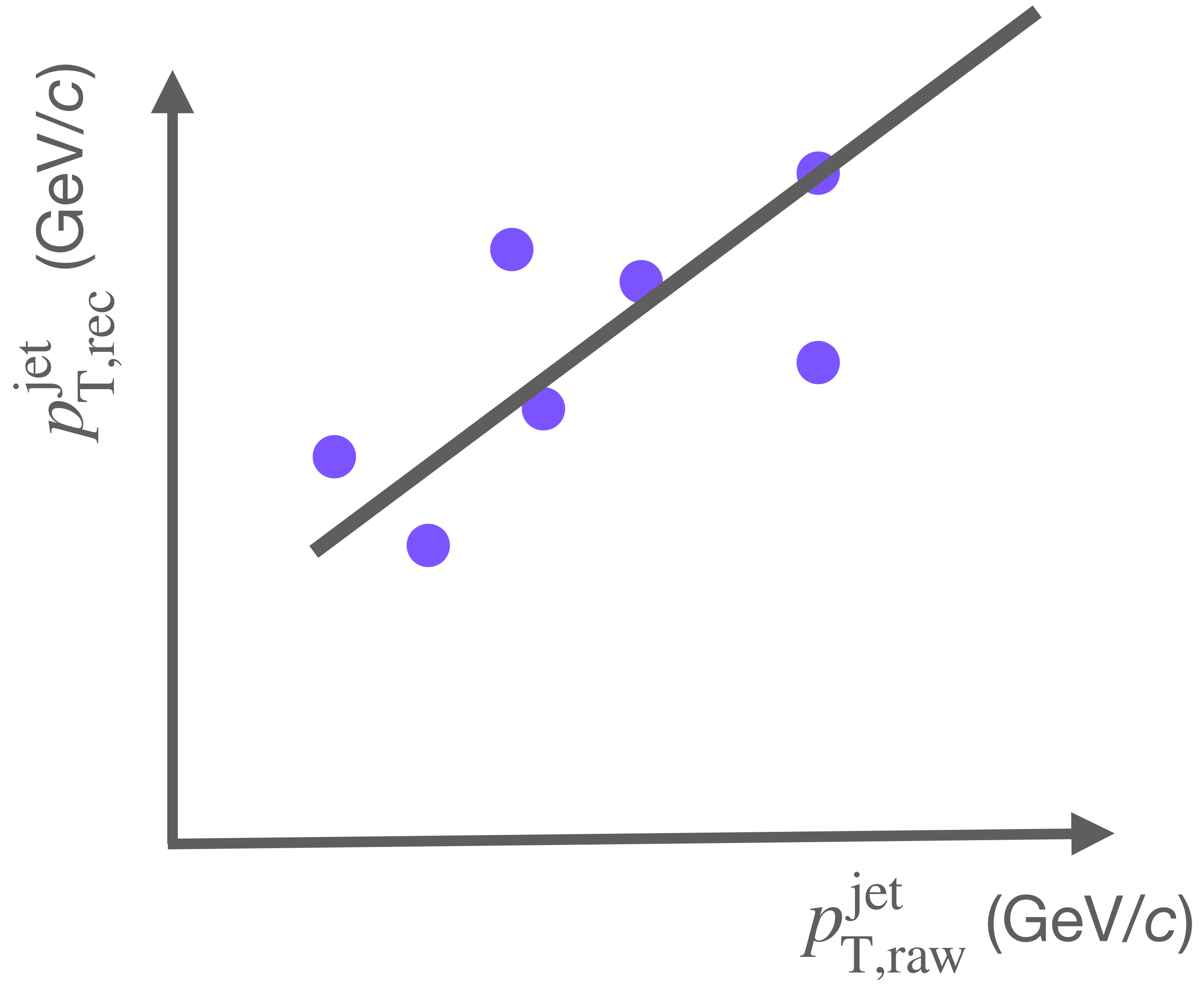
Indirect training → generative network never sees the true distribution!



Intro to Linear Regression

Linear regression predicts the value of a **dependent variable** based on a given **independent variable** (feature x1 with a given weight w1).

$$y = b + w_1 x_1$$



The example at the right is a simplified view in reality we have multiple features each having a separate weight.

$$y = b + w_1 x_1 + w_2 x_2 + w_3 x_3 \ldots\ldots$$

Training determines the optimal weight for each feature.

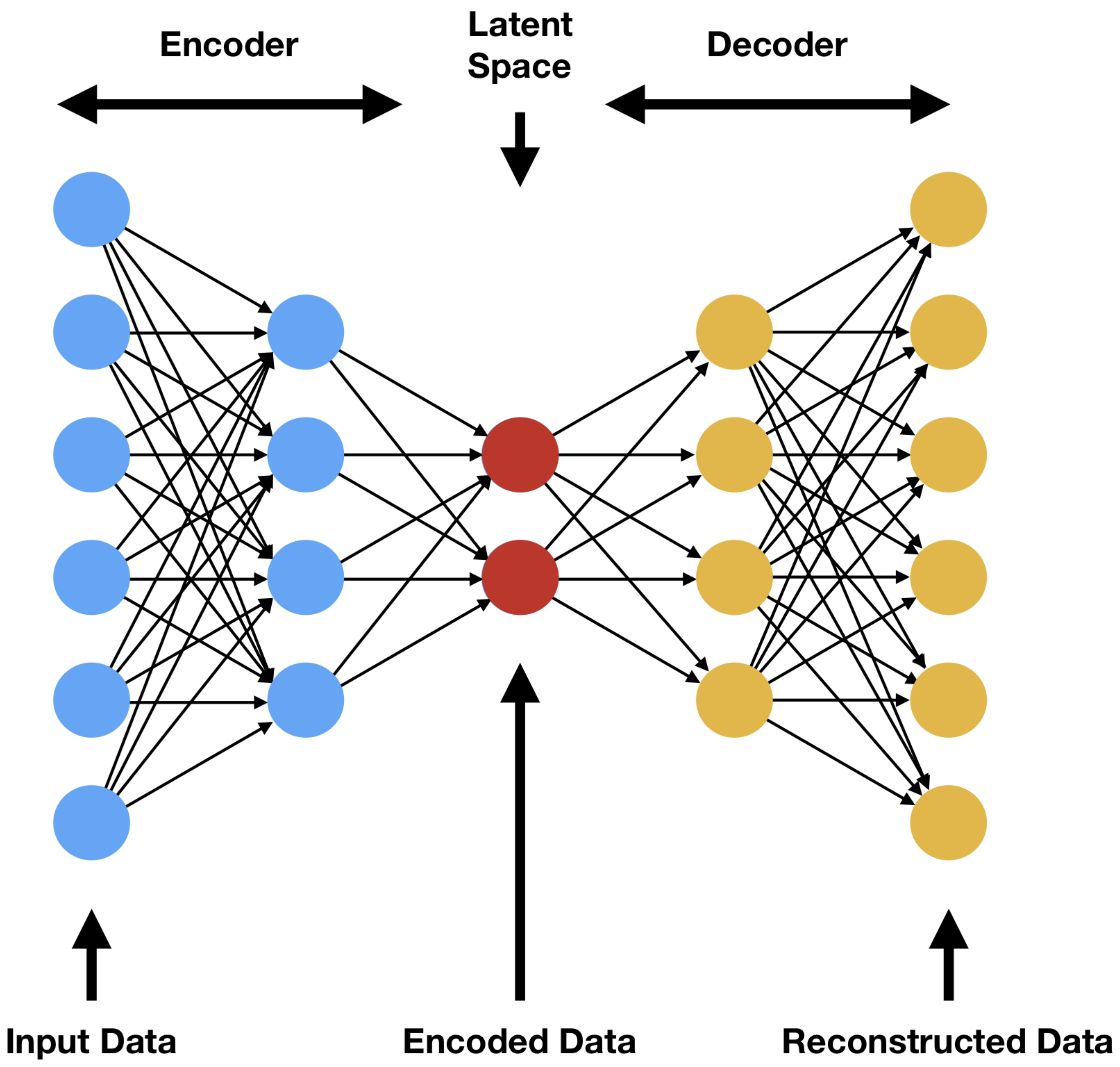
Auto-Encoders

Simple task: NN architecture trained to copy inputs to outputs!

Encoder takes the input and dramatically reduces its complexity via a NN.

Decoder takes the encoded data and reconstructs outputs like the data.

Does not require labeled data as input!

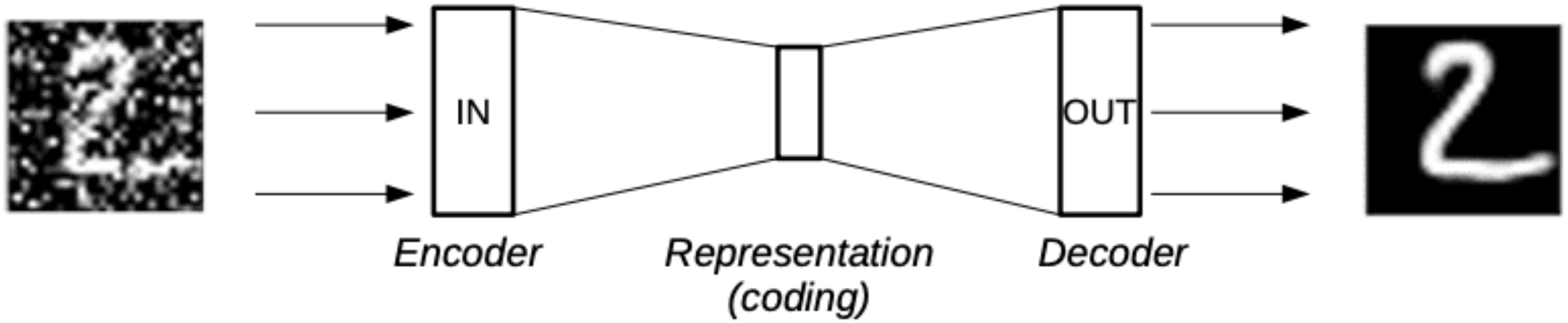


https://www.compthree.com/blog/autoencoder/

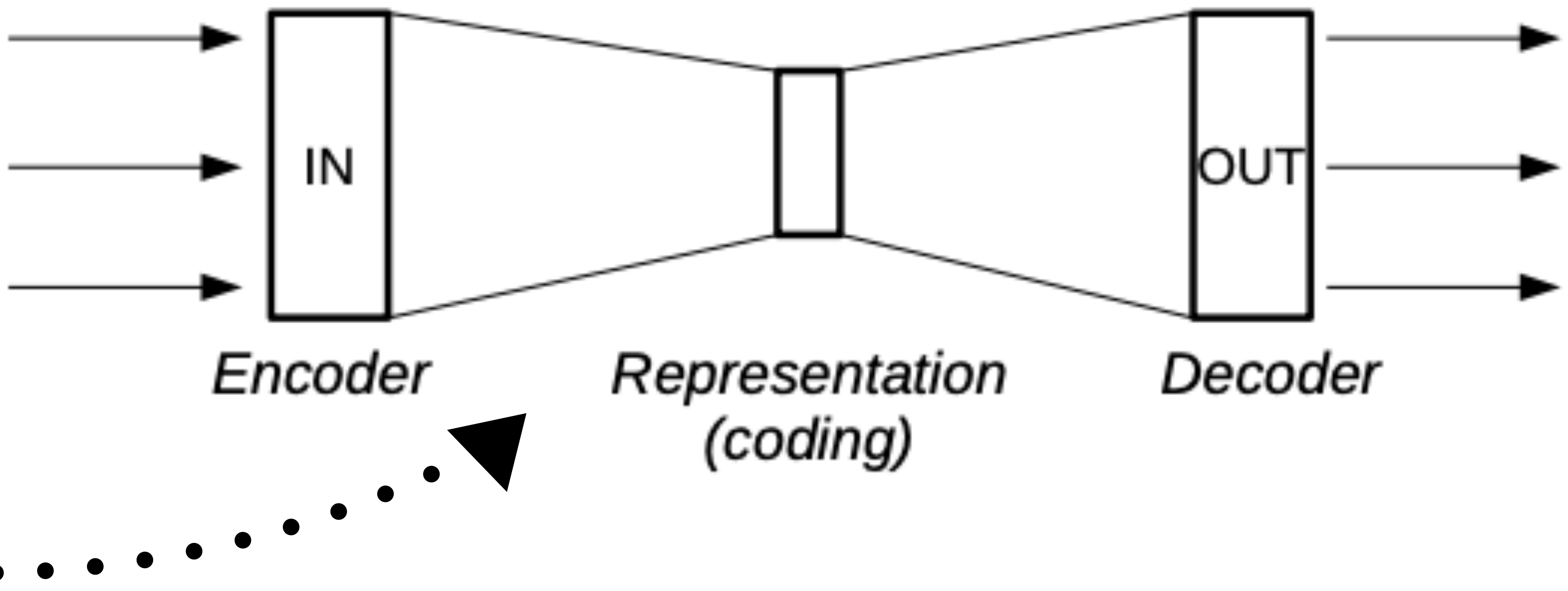
Uses of Auto-Encoders

Used to learn efficient representations of some input data.

1. De-noising inputs



2. Unsupervised learning



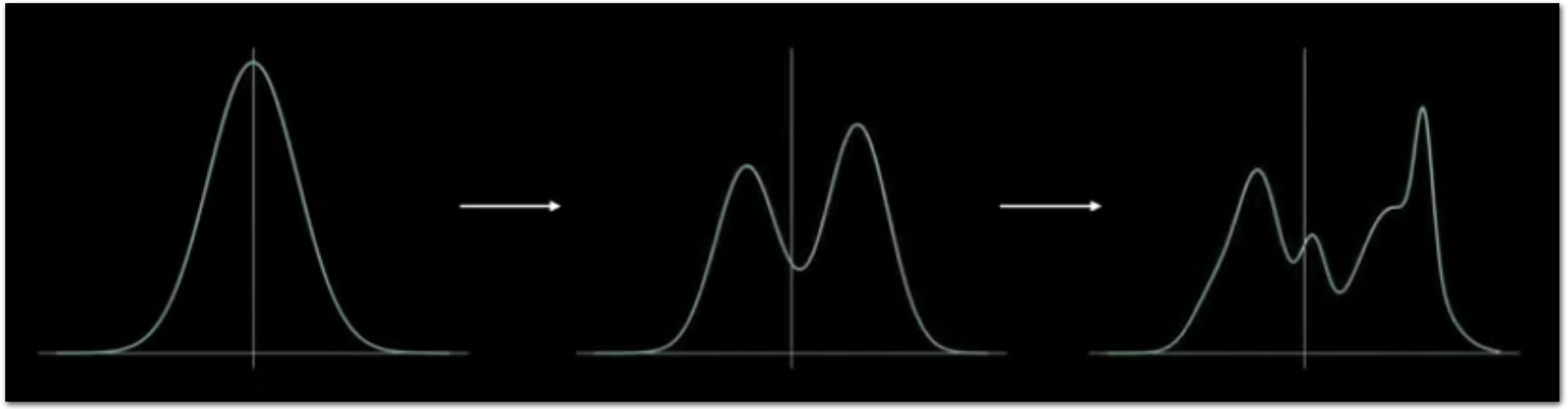
Sort items into classes here

3. Anomaly detections: If you fail to reconstruct data in the decoding step you have an anomaly!

Normalizing Flows

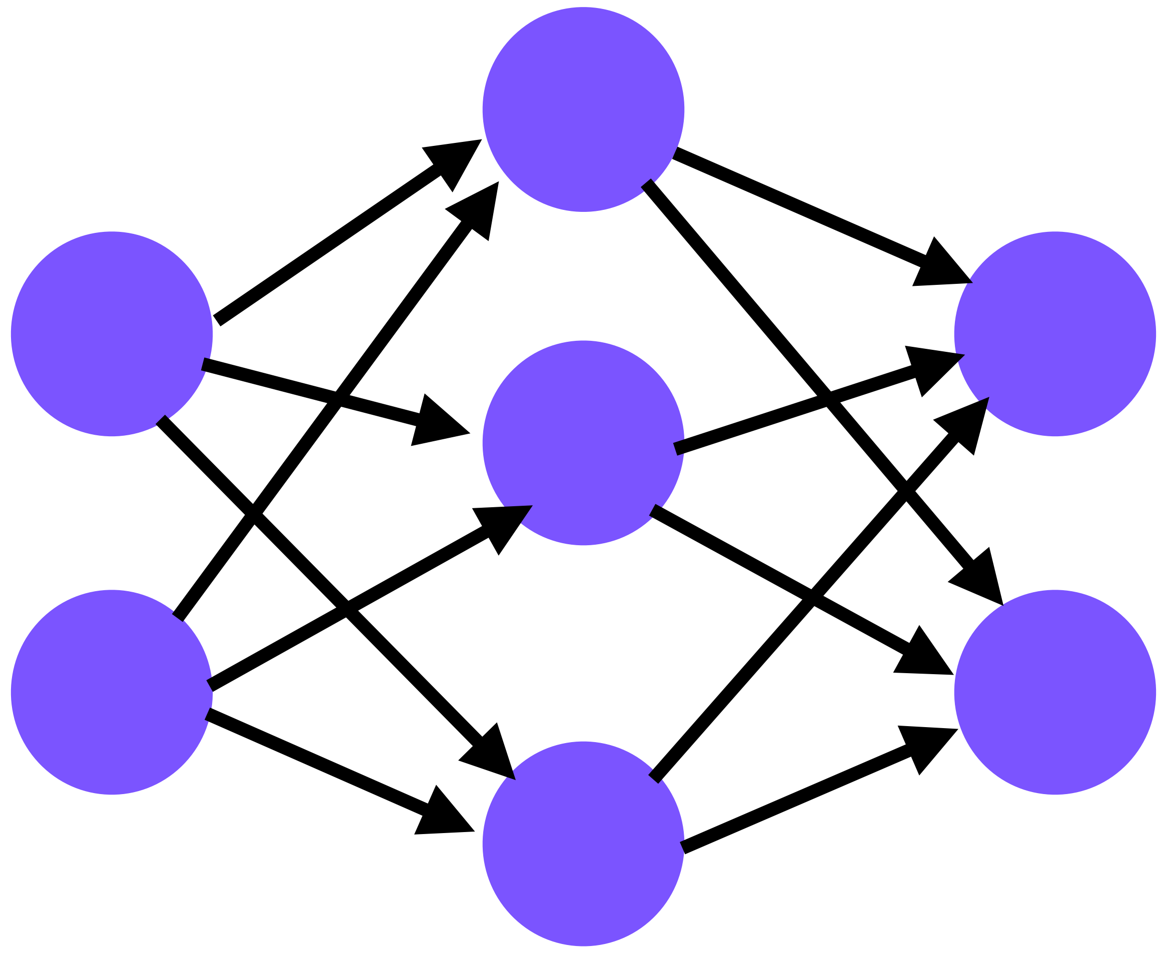
Generative modeling tool used to build complex probability distributions by transforming simple ones.

Make individual transformations between probability distributions invertible so the overall transformation is also invertible.

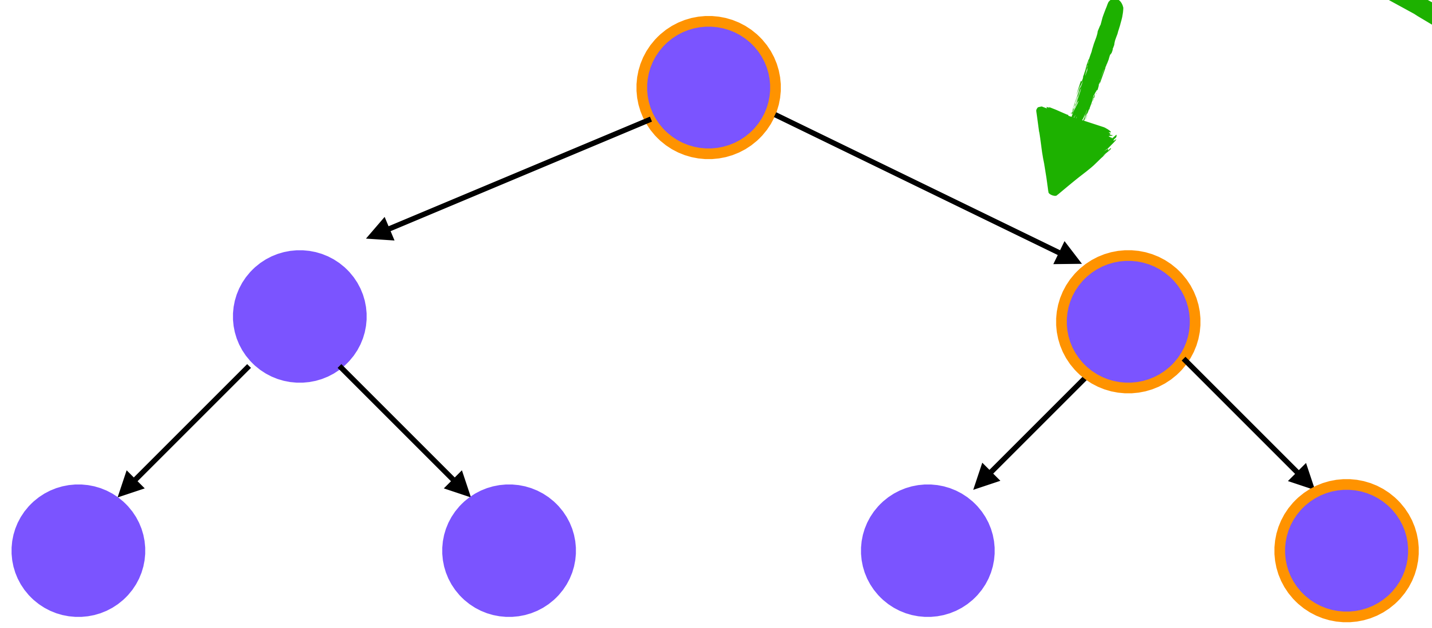


Can use probability distributions to sample likelihood distributions!

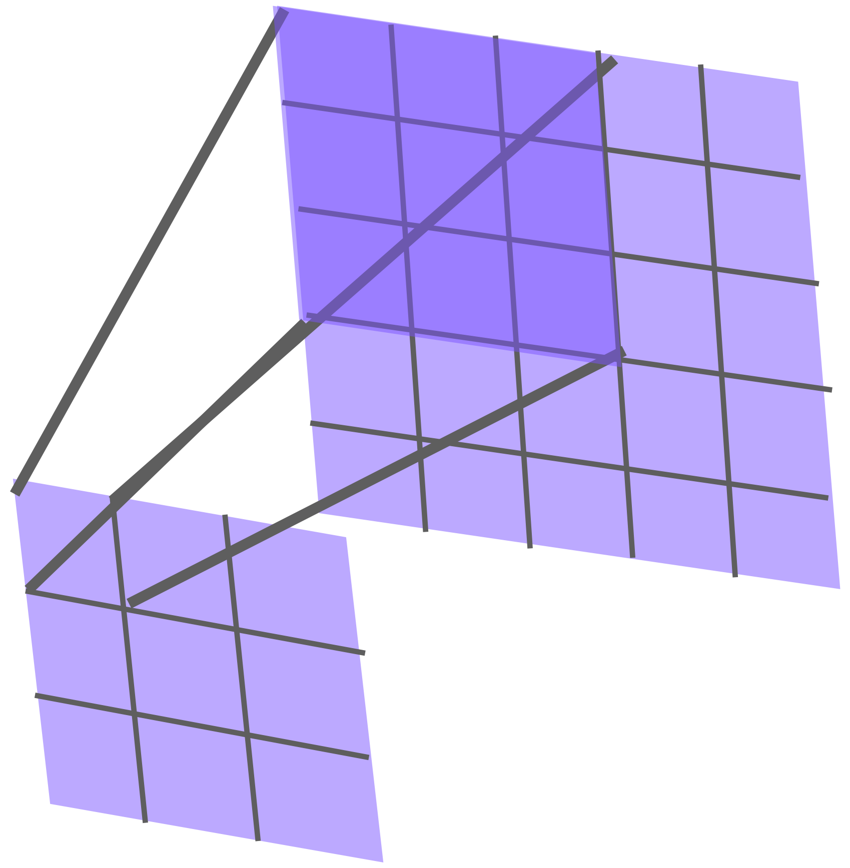
Different algorithms for different problems!



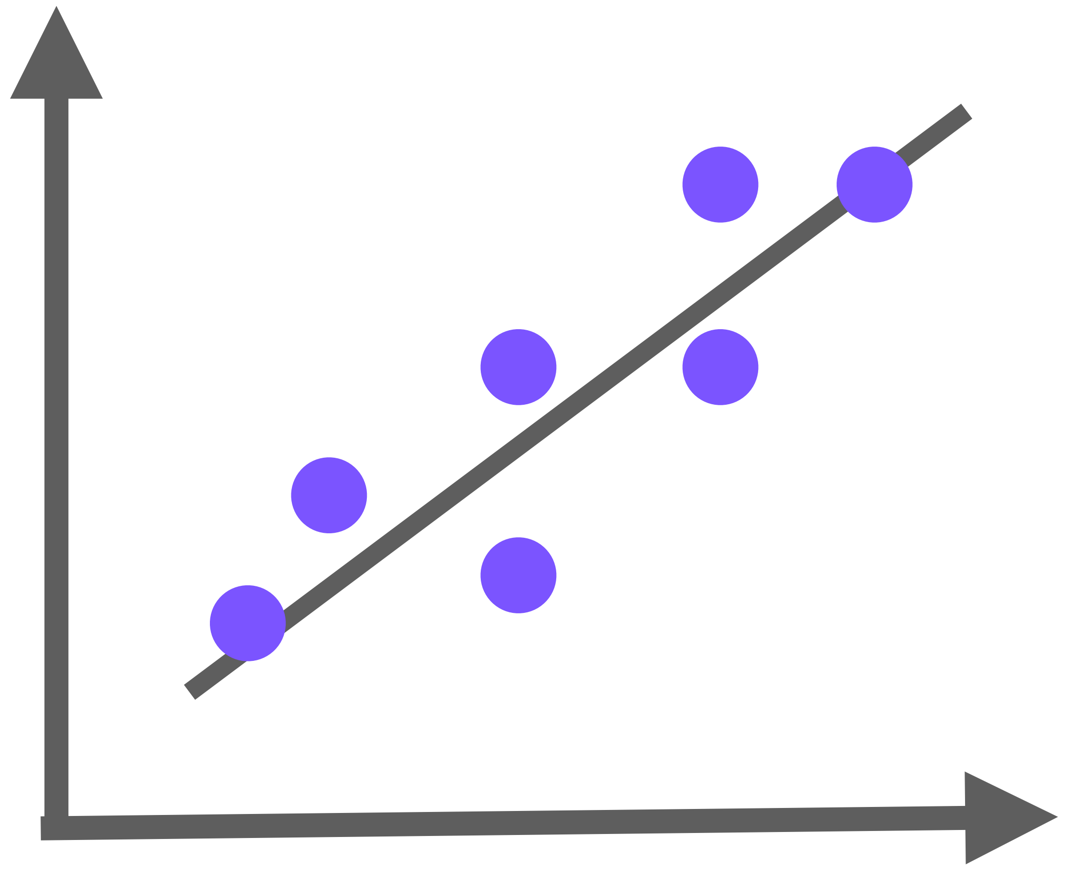
(Shallow or Deep) Neural Networks → *Great for making predictions!*



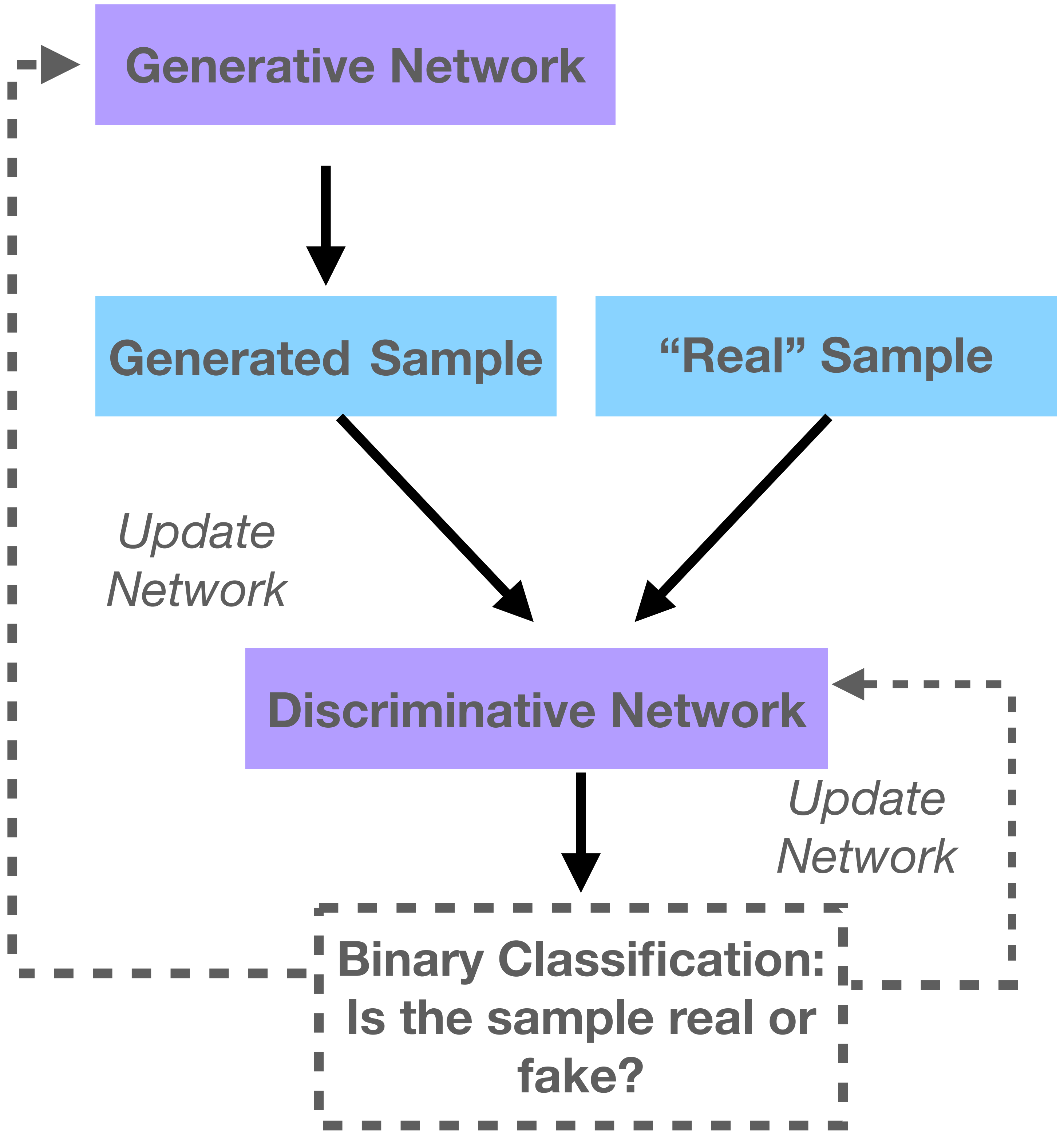
Random Forest (Decision Trees)



Convolutional Neural Networks (CNNs)→ *Great for image processing!*



Linear Regression

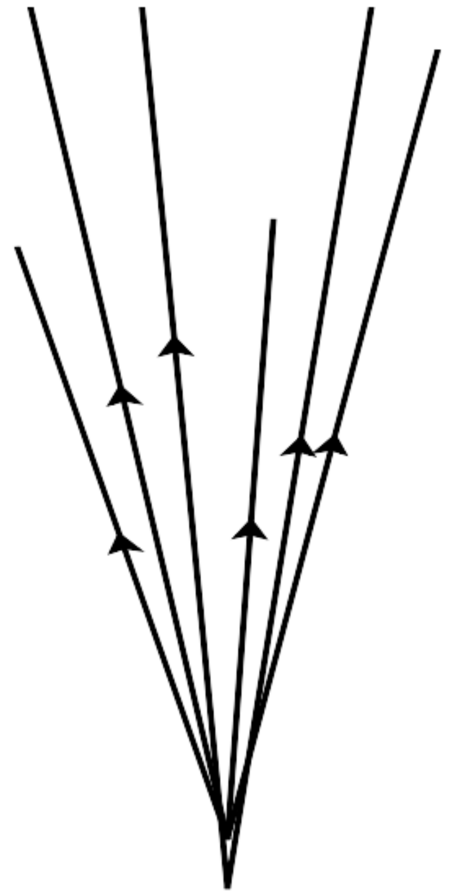


Generative Adversarial Networks (GANs) → *Powerful tool for generating samples!*

Jet VLAD architecture

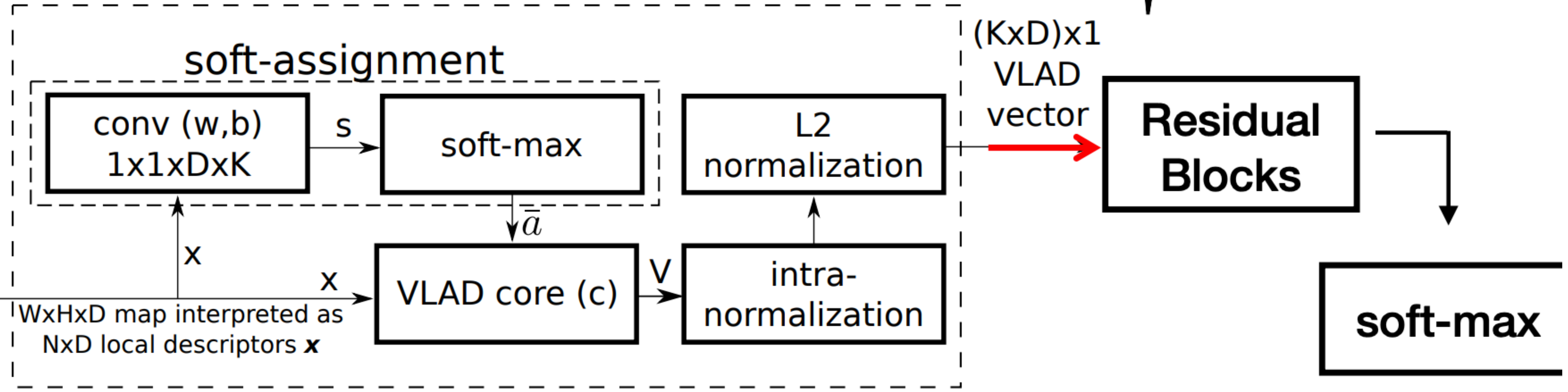
Machine Learning   JetVLAD @ RHIC

$$\mathcal{J} = \{(p_{T,i}, \eta_i, \phi_i, \ldots)\}_{i=1}^{n} \,,$$



$$V(:,k) = \sum_{i=1}^{N} \frac{e^{w_k^T x_i + b_k}}{\sum_{k'} e^{w_{k'}^T x_i + b_{k'}}} (x_i - c_k)$$

NetVLAD layer



Vector of Locally Aggregated Descriptors

Ponimatkin, et. al
JINST (2021) 2005.01842

Total of 111608 trainable parameters

D - Depth
K - # Clusters

Slide from Ragahav Kunnawalkam Elayavalli

Uses of ML for High Energy Physics

The LHC Olympics 2020

A Community Challenge for Anomaly Detection in High Energy Physics

Anomaly detection (LHC Olympics, arXiv:2101.08320)

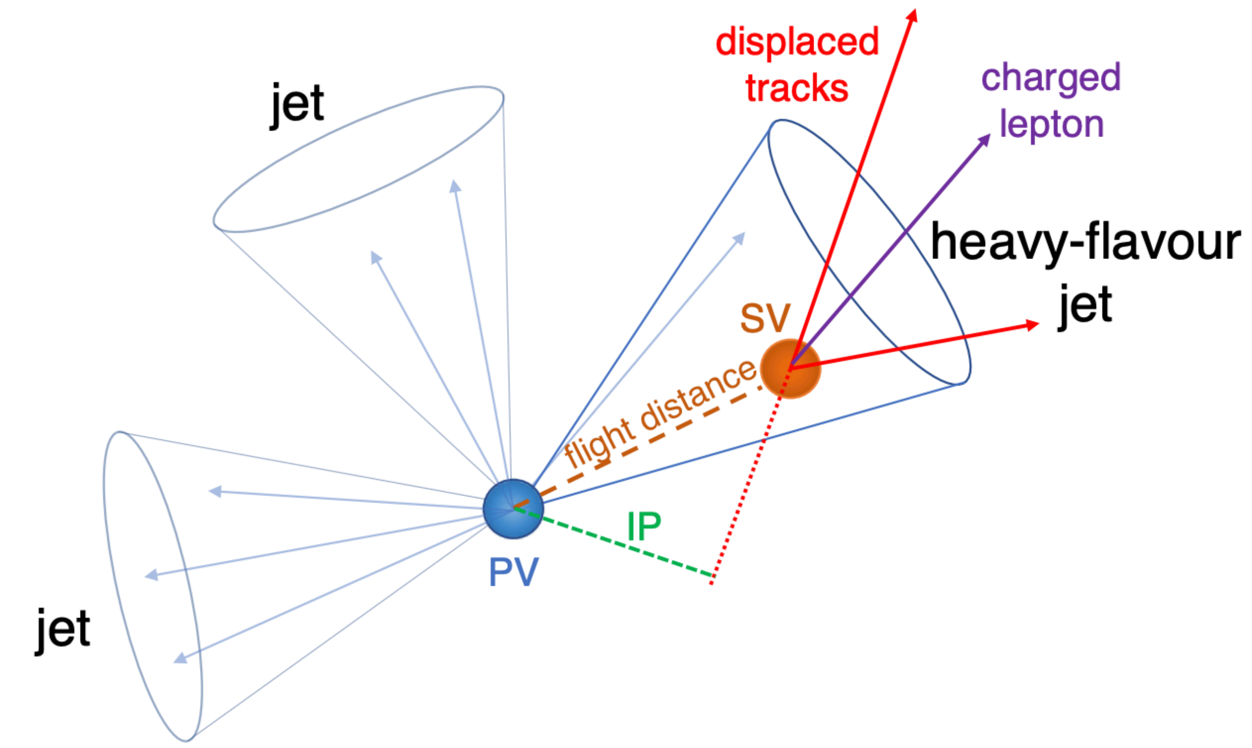
Many broad categories of applications!

Triggering Systematic Uncertainty Reduction

Anomaly Detection Object Classification

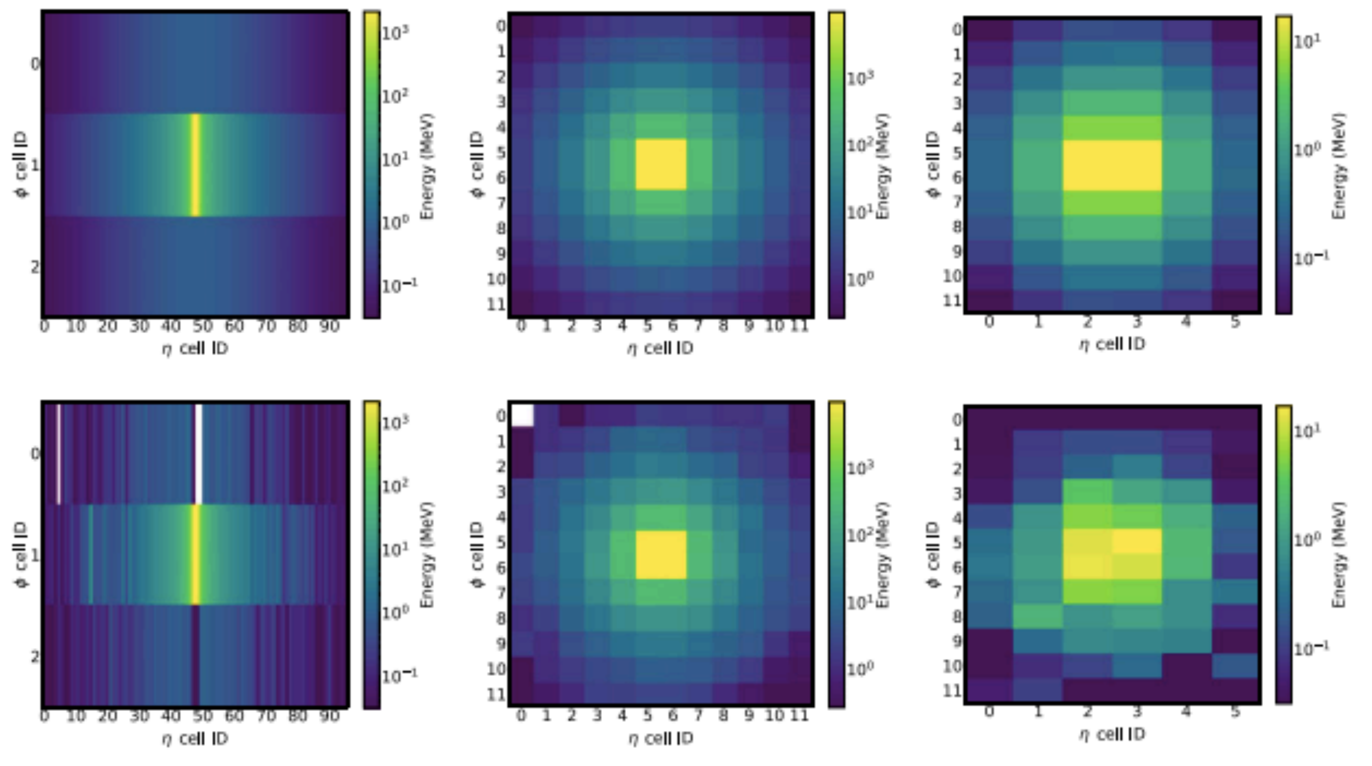
Data Quality Assessment Detector Simulation

New Physics Searches



Object Classification arXiv: 1712.07158

What about heavy-ion physics specifically?



Geant

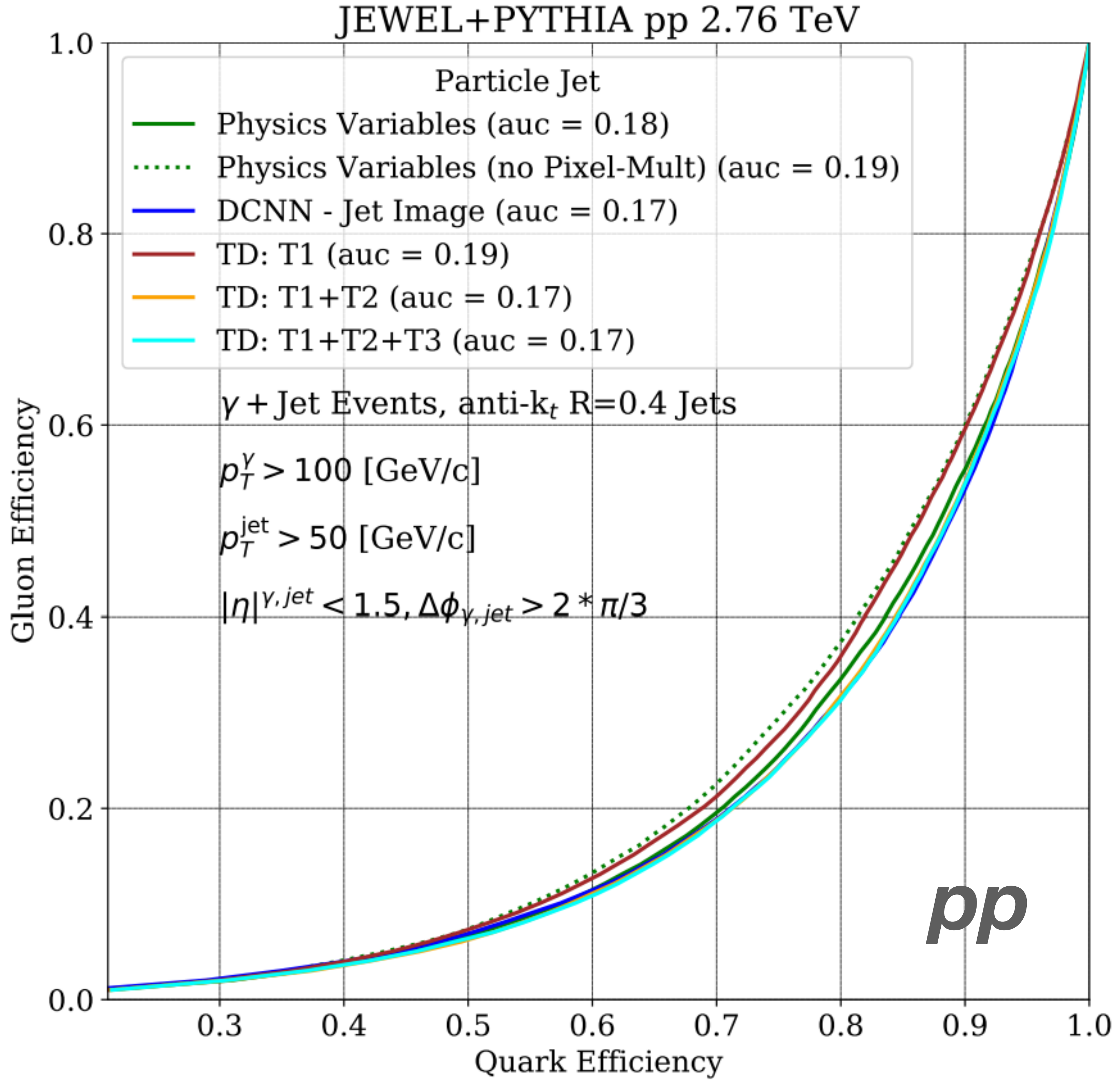
CaloGAN

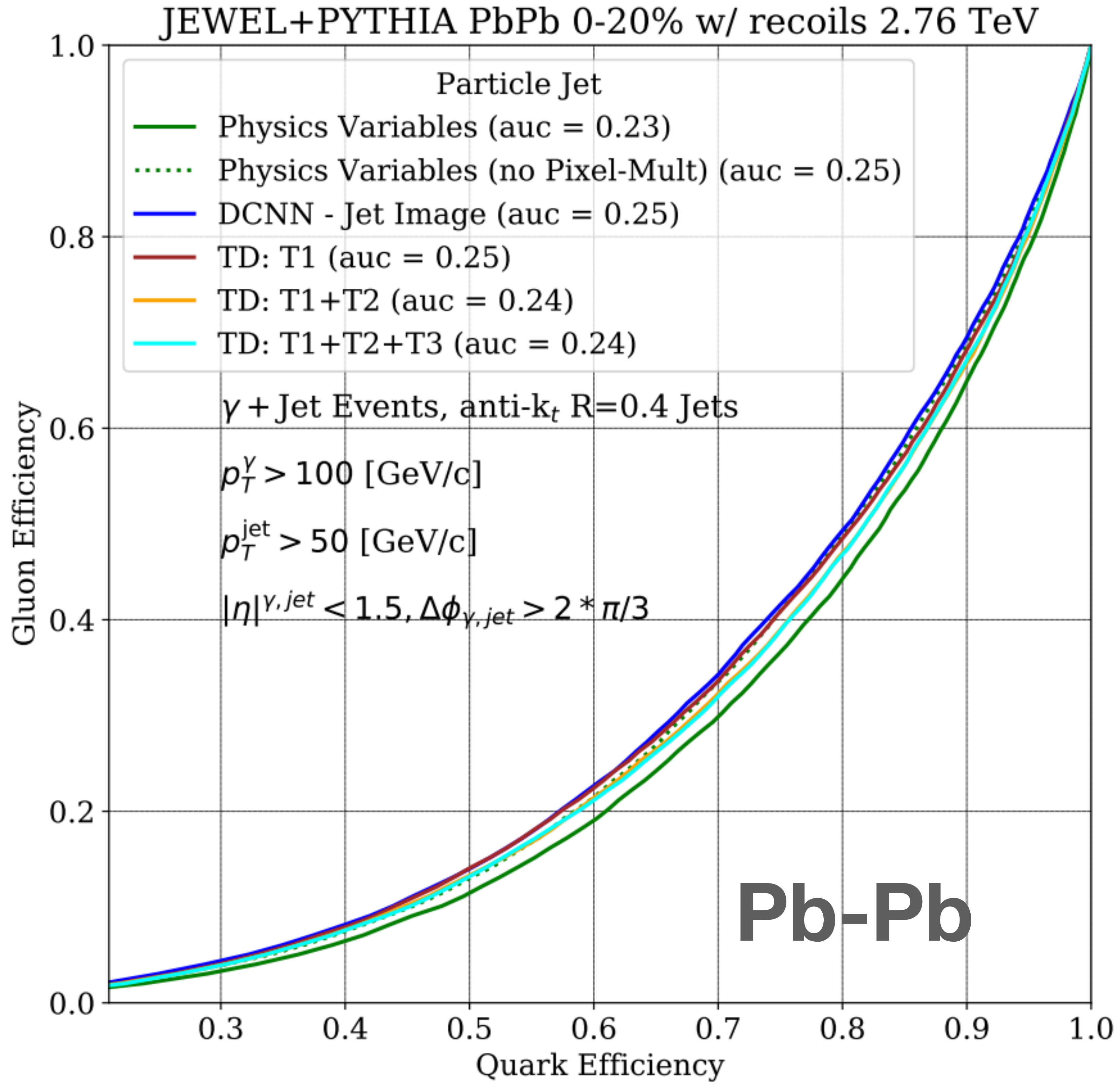
Detector Simulation

(CaloGAN, PRL 120, 042003 (2018))

Quark vs. Gluon Jets with ML

Y. Chien, R. Elayavalli: 1803.3589





Lower the curve, the better the performance.

All methods explored tend to perform consistently, indicating that they may be picking up on similar features.

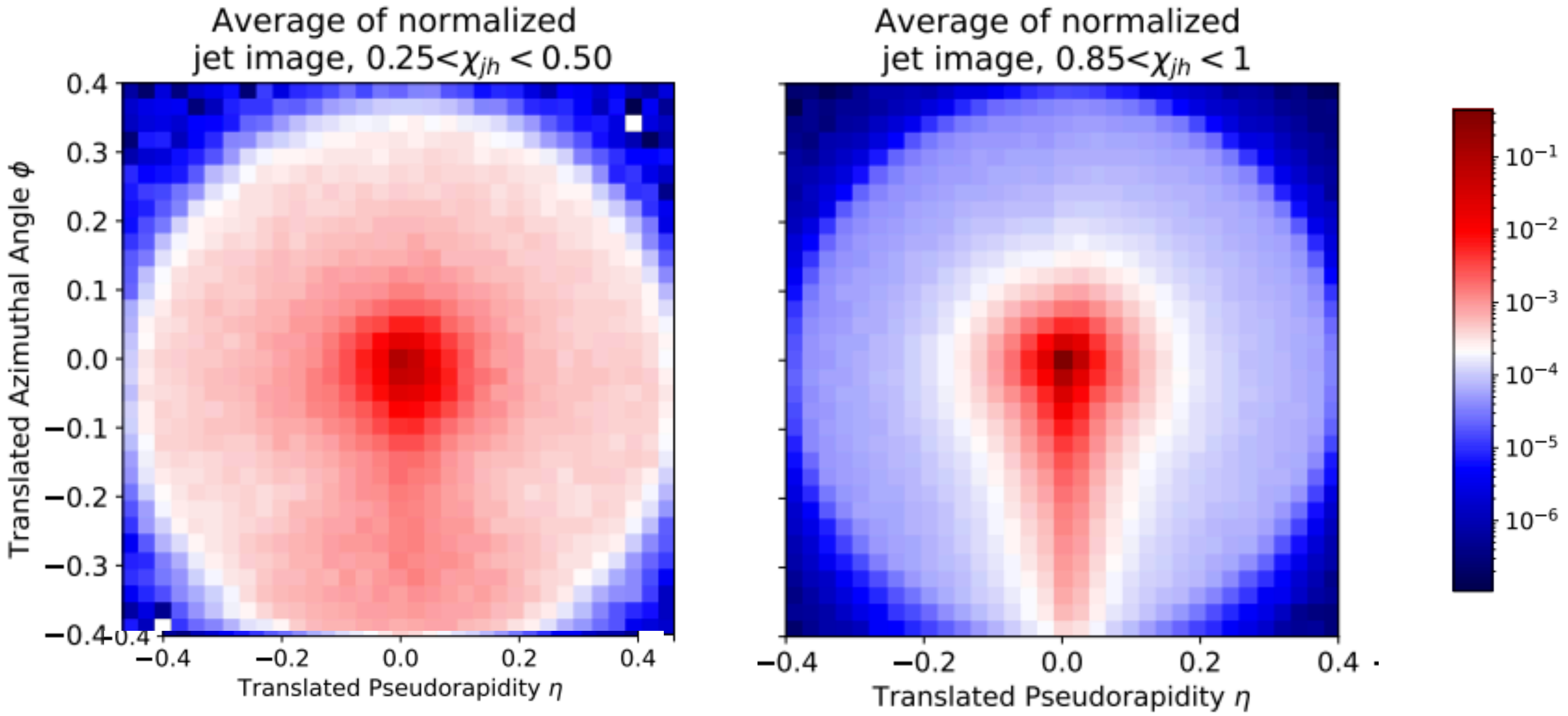
The performance worsens for Pb—Pb, due to the large UE.

Quark and gluon discrimination is a difficult and ongoing effort in HIs!

Future: Apply these methods to data in pp and Pb—Pb!

Deep Learning Jet Modifications

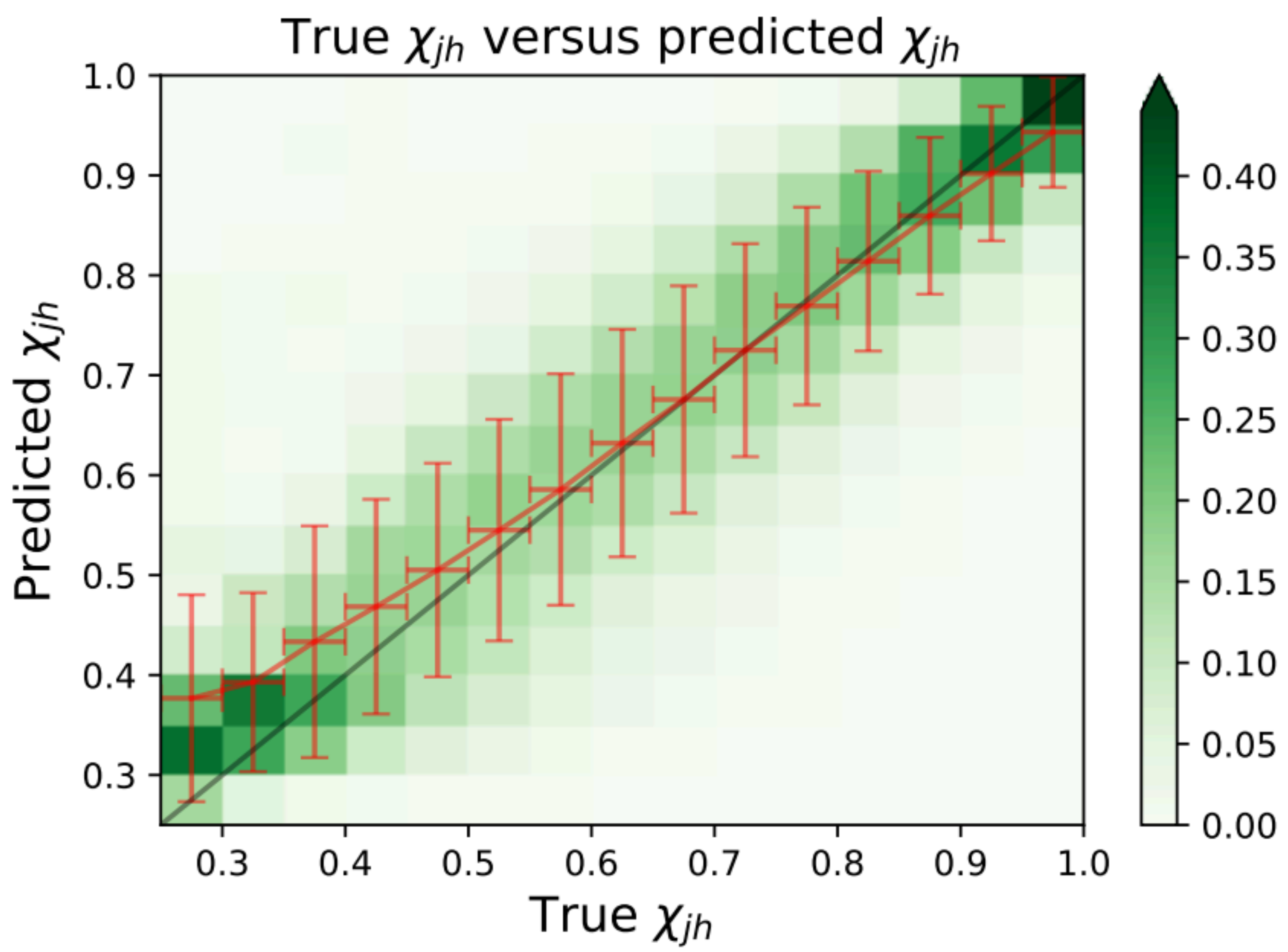
Use supervised learning on jet images with a CNN to perform the regression task of predicting the energy loss ratio in HI collisions (hybrid model).



quenched unquenched

$$\chi_{jh} = \frac{E_f^h}{E_i^h}$$

Y. Du, D. Pablos, K. Tywoniuk: 2101.07797



Shows good performance!

Very useful to separate and study quenched vs. unquenched jets as well as extracting the initial energy of the jet.

Future: Apply these methods to different models & variables, improve performance.

Far future: Apply these methods to data!