

Honors Report Spring 2025

Study of Machine Learning models in the field of Particle Physics.



INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY

H Y D E R A B A D

Abhiram Tilak - 2022113011

**International Institute of Information Technology, Hyderabad*

Abstract: This report goes over a detailed classification of the different types of machine learning models used in the field of particle physics.

Note that the purpose of this study is not to perform a deep dive on the architectural specifications of each model, most of the discussion focuses on the applications of the models in the field of particle physics. We present a brief overview of the different types of models used in the field of particle physics.

Contents

1	Introduction	2
1.1	Types of Tasks in Particle Physics	3
2	Different types of models in HEP	5
2.1	Generative Models	5
2.1.1	CaloGAN[1]	7
2.1.2	Variational Autoencoders	7
2.2	Discriminative Models	7
2.2.1	ParticleNet[2]	8
2.3	Unsupervised Learning Models	10
2.4	Reinforcement[3] Learning in HEP	10
3	Conclusion	12

List of Figures

1	An example of jet tagging, where jets are classified based on the type of fundamental particle that produced them.	4
2	In ML for HEP, anomaly detection is used to identify events that deviate from expected patterns, potentially pointing to new physics beyond the Standard Model.	5
3	A graphical representation of a probabilistic model.	6
4	architecture of a variational autoencoder (VAE) for generating realistic and diverse jet images.	7
5	Architecture of ParticleNet	9
6	architecture of EdgeConv block in ParticleNet	9

List of Tables

1	Classification of selected machine learning and deep learning models in HEP by learning paradigm.	11
---	---	----

Listings

1. Introduction

ML methods are designed to harness large amounts of data in order to infer new information, making them naturally suited to various data processing tasks in Particle Physics, from the moment

a particle hits the detector all the way to the final published results. The examples that follow give a representative idea of how extensively this technology has pervaded the experiment, but merely scratch the surface of the full picture and potential of ML in HEP.

The application of machine learning techniques is gaining momentum in the research field of high energy physics (HEP) and astroparticle physics. The experiments at large hadron collider (LHC) as well as several other collider-based and astroparticle experiments are accumulating large amounts of data for the precision measurement of parameters of the Standard Model of particle physics and to search for existence of beyond Standard Model physics at higher energy scale for which there are compelling theoretical and experimental reasons.

In the future, the high luminosity LHC is expected to deliver ten times more data than what is available till date. Already, remarkable progress has been achieved in the application of machine learning in HEP, in terms of developing event classification, object identification, and estimation strategies. ML methods are expected to be heavily employed in future data analysis.[4]

1.1. Types of Tasks in Particle Physics

Before diving into different ML models, it is important to understand the types of tasks that are performed in particle physics to understand why machine learning was even considered in the first place.

We can divide these tasks into two broad categories:

- **Classification:** Categorizing events or particles into predefined classes.
 - *Jet Tagging:* Classifying jets (collimated sprays of particles) based on the type of fundamental particle (e.g., quark, gluon, or heavy quarks like b or top quarks) that produced them. This is crucial for identifying decay products and understanding underlying physical processes.
 - *Particle Identification:* Distinguishing between particle types (e.g., electrons, muons, protons, pions) using their detector signatures. This is essential for accurate event reconstruction.
 - *Signal vs. Background Discrimination:* Detecting rare signal events amidst a larger background. This is often framed as a binary classification problem.
 - *Event Filtering (Trigger Systems):* Real-time classification in detector trigger systems to decide which events to retain for analysis due to high collision rates.
- **Regression:** Predicting continuous values from data.

- *Energy and Momentum Regression*: Estimating particle or jet energy and momentum from detector signals, which often exhibit complex, non-linear responses.
- *Vertex Position Estimation*: Determining the spatial origin of particle interactions within the detector.
- *Tracking*: Reconstructing trajectories of charged particles. While tracking involves pattern recognition, refining trajectory parameters is treated as a regression task.

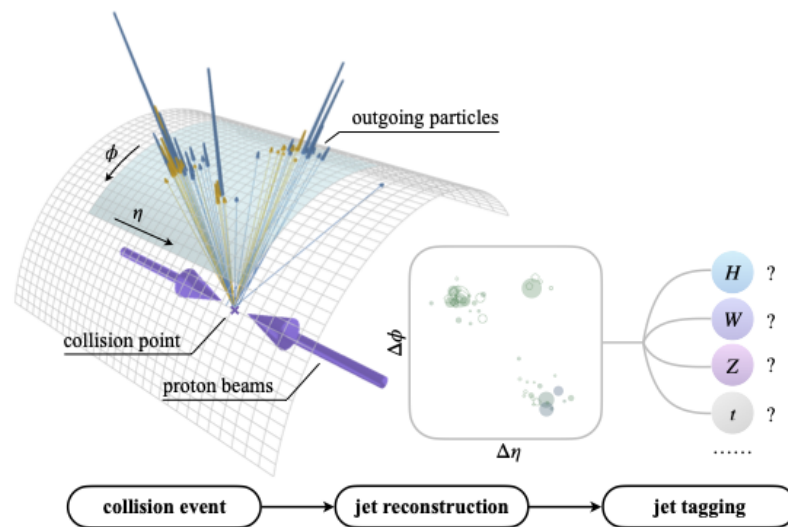


Fig. 1: An example of jet tagging, where jets are classified based on the type of fundamental particle that produced them.

There are also methods don't fit neatly into these categories, such as:

- **Anomaly Detection**: Identifying events that deviate from expected patterns, potentially pointing to new physics beyond the Standard Model.
- **Simulation**: Generating realistic synthetic data for analysis and comparison with theoretical models.
- **Unfolding**: Correcting for detector effects to recover true distributions of physical quantities. Machine learning models are used to invert the detector response.

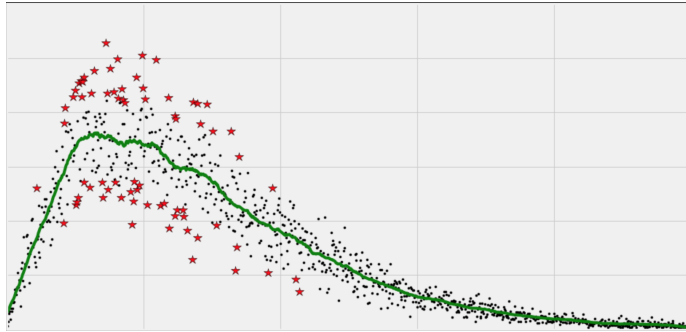


Fig. 2: In ML for HEP, anomaly detection is used to identify events that deviate from expected patterns, potentially pointing to new physics beyond the Standard Model.

2. Different types of models in HEP

Based on the architecture used by the models, we can divide models into various categories.¹

2.1. *Generative Models*

Generative models are particularly valuable in HEP for tasks involving the creation of realistic simulated data or modeling complex data distributions.

Generative models focus on learning underlying probability distributions to generate new, synthetic samples that resemble real data. These models are particularly valuable in HEP for simulation and data augmentation tasks.

¹**Note:** Each of these models can solve various different tasks in HEP, and the same task can be solved using different models. The choice of model depends on the specific problem, data characteristics, and computational resources available. For instance if we take an example from the NLP field can be adapted to solve a different task. For example, BERT, which is primarily used for text classification, can also be adapted for tasks like named entity recognition or question answering. Similarly, models like GPT-3, originally designed for text generation, can be fine-tuned for various NLP tasks such as summarization or translation.

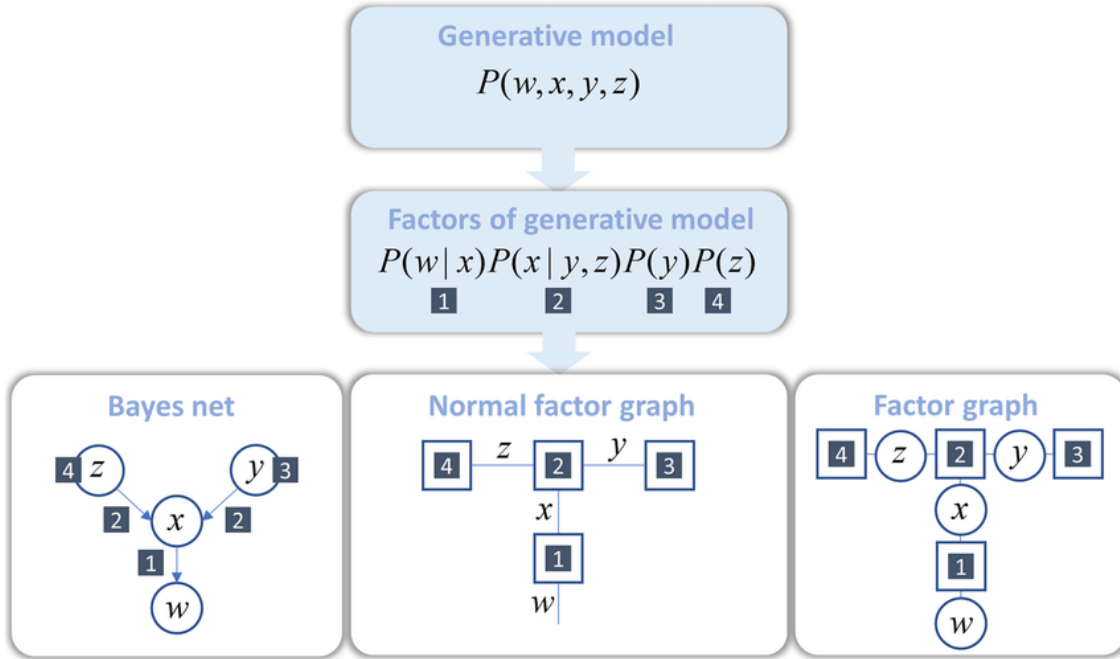


Fig. 3: A graphical representation of a probabilistic model.

- **Fast Detector Simulation:** Traditional simulation methods like GEANT4[5] are computationally expensive. Generative models such as GANs and VAEs are being developed to produce synthetic detector responses that are significantly faster while maintaining fidelity to real detector characteristics. This is essential for generating large simulated datasets, especially as LHC luminosity increases.
- **Event Generation:** Generative models are being explored to model and generate entire events or specific components, with the goal of accelerating parts of the simulation chain.
- **Modeling Complex Distributions:** These models can learn the probability distributions underlying high-dimensional collision data, aiding in understanding event structures and supporting tasks like anomaly detection.
- **Data Augmentation:** Creating synthetic but realistic variations of existing data to expand training datasets, particularly for rare processes.

Some popular examples of models in this category include:

2.1.1. CaloGAN[1]

CaloGAN represents a breakthrough application of Generative Adversarial Networks (GANs) for simulating electromagnetic showers in multi-layer calorimeters. Developed to address the computational bottleneck in traditional Monte Carlo simulation methods, CaloGAN can generate synthetic calorimeter responses with remarkable speed improvements. What differentiates CaloGAN from other generative approaches is its physics-aware architecture that preserves essential calorimeter shower properties while maintaining the enormous speed advantage. This model demonstrates how deep learning techniques can significantly accelerate one of the most computationally expensive steps in HEP simulation pipelines without sacrificing physical accuracy.

2.1.2. Variational Autoencoders

Another important generative approach in HEP employs Variational Autoencoders (VAEs). A notable implementation described in a 2020 paper introduces a specialized VAE architecture for generating realistic and diverse jet images in high energy physics events. The latent space characteristics of VAEs provide an interpretable representation that can be manipulated to generate specific physics scenarios, offering advantages for tasks like jet tagging evaluation and uncertainty quantification that may be more challenging with other generative approaches.

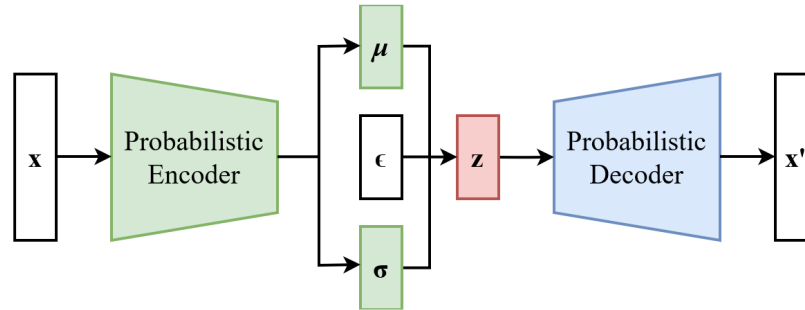


Fig. 4: architecture of a variational autoencoder (VAE) for generating realistic and diverse jet images.

2.2. Discriminative Models

Discriminative models are widely used for tasks where the goal is to differentiate between classes or predict values, typically under supervised learning.

Discriminative models excel at classification and regression tasks, learning boundaries between different classes of data-capabilities that are essential for many core HEP analysis requirements.

- **Jet Tagging:** Classification of jets based on particle content and energy distribution to determine origin (e.g., b -jet, top-jet, quark-jet, gluon-jet), using models like deep neural networks, BDTs, and SVMs.
- **Particle Identification (PID):** Classifying tracks or energy deposits to identify particles such as electrons, muons, photons, and protons—critical for event reconstruction.
- **Signal vs. Background Classification:** Classifying events to separate Standard Model backgrounds from potential new physics signals—a core task in discovery physics.
- **Event Filtering (Triggers):** Discriminative models are used in trigger systems to make real-time decisions about which events to keep for offline analysis.
- **Regression for Physics Object Properties:** Predicting properties such as jet energy or vertex positions from raw detector data using regression models.

2.2.1. ParticleNet[2]

ParticleNet represents a state-of-the-art discriminative model specifically designed for jet tagging tasks in particle physics. The model employs a graph neural network architecture that treats jets as "particle clouds"—collections of individual particles with their associated properties.

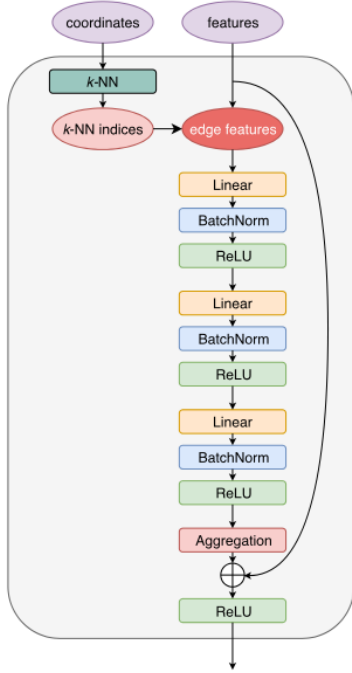
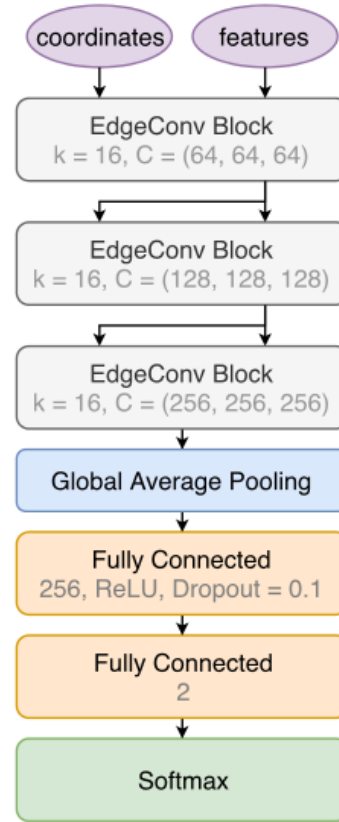


Fig. 5: The EdgeConv block in ParticleNet operates on particle clouds by dynamically building a graph for each particle based on its k -nearest neighbors. It computes edge features representing the relationship between a particle and its neighbors, which are then processed by a shared MLP. An aggregation step combines these features to create an enriched feature vector for each particle, allowing the network to learn local structures in the particle cloud in a permutation-invariant manner.



(a) ParticleNet

Fig. 6: The architecture of the ParticleNet network. The ParticleNet architecture consists of three EdgeConv blocks. After the EdgeConv blocks, a channel-wise global average pooling operation is applied to aggregate the learned features over all particles in the cloud. This is followed by a fully connected layer with 256 units and the ReLU activation, and a dropout layer. A fully connected layer with two units, followed by a softmax function, is used to generate the output for the binary classification task.

What sets ParticleNet apart from alternative approaches is its significantly improved performance combined with computational efficiency. When compared to image-based deep neural networks like ResNeXt-50 or other particle-based networks like P-CNN and PFN, ParticleNet achieves superior performance on both top tagging and quark-gluon tagging benchmarks. Remarkably, it accomplishes this with substantially fewer parameters—4 times fewer than ResNeXt-50 for

the standard version and 56 times fewer for the lightweight version. This efficiency makes ParticleNet particularly valuable for applications with stringent latency and memory constraints, such as online event processing in HEP experiments.

2.3. *Unsupervised Learning Models*

Unsupervised learning is applied when labels are not available, with the goal of discovering patterns or anomalies.

Unsupervised learning models discover patterns in data without labeled examples, making them valuable for anomaly detection and dimensionality reduction in HEP contexts.

- **Anomaly Detection:** Detecting events that deviate from expected Standard Model patterns, potentially signaling new physics. Techniques include autoencoders (which detect anomalies via reconstruction error) and clustering approaches.
- **Clustering:** Grouping similar data points (e.g., calorimeter deposits or tracks). ML-based clustering can enhance or replace traditional algorithms in tasks such as vertex or shower reconstruction.
- **Dimensionality Reduction:** Reducing input feature space for visualization or preprocessing using techniques like PCA and t-SNE, while retaining essential data characteristics.

2.4. *Reinforcement[3] Learning in HEP*

Although less commonly used for direct data analysis, reinforcement learning is increasingly applied to system optimization.

- **Particle Accelerator Control:** Tuning accelerator parameters to optimize performance metrics like beam intensity and stability via interaction with control systems.
- **Detector Operation Optimization:** Potential use in real-time detector configuration to enhance data acquisition and quality.
- **Event Reconstruction Algorithms:** Exploring sequential decision-making with reinforcement learning in reconstruction tasks, such as deciding how to associate detector hits into tracks.

Table 1: Classification of selected machine learning and deep learning models in HEP by learning paradigm.

Model Name	Category	Primary Application(s)	Notable Features / Comments
CaloGAN[1]	Generative	Calorimeter simulation (3D EM showers)	GAN-based generative model for fast simulation of electromagnetic showers ² .
Normalizing Flow (NF)	Generative	End-to-end event simulation	Flow-based invertible generative model with exact likelihood; used for full-chain event generation ³ .
Variational Autoencoder (VAE)	Unsupervised	Anomaly detection (new physics searches)	Latent-variable generative model for unsupervised anomaly detection in events ⁴ .
Autoencoder (AE)	Unsupervised	Anomaly detection (trigger filtering)	Compresses events and flags unusual events via high reconstruction error ⁵ .
Convolutional Neural Network (CNN)	Discriminative	Calorimeter imaging / trigger	Convolutional classifier for calorimeter data (pattern recognition); e.g. identifying displaced particle showers ⁶ .
DeepJet	Discriminative	Jet flavor tagging (b/c jets)	Deep neural network using full jet constituent information; improves heavy-flavor tagging performance ⁷ .
ParticleNet[2]	Discriminative (Graph-based)	Jet tagging (quark/gluon, top)	Dynamic graph CNN on particle clouds; permutation-invariant jet classifier ⁸ .
LorentzNet[6]	Discriminative (Graph-based)	Jet tagging	Lorentz-equivariant graph neural network for jets; uses Minkowski dot-product attention ⁹ .
GNN tracking pipeline[7]	Graph-based (Supervised)	Charged particle track reconstruction	Graph neural network (message-passing) linking detector hits into tracks ¹⁰ .

3. Conclusion

These diverse machine learning approaches demonstrate how computational techniques are transforming high-energy physics research across multiple fronts. Generative models like CaloGAN and VAEs are dramatically accelerating simulation tasks. Discriminative models like ParticleNet are enhancing our ability to identify specific particle signatures. Unsupervised approaches including quantum autoencoders enable model-agnostic searches for new physics. Reinforcement learning techniques are optimizing the operation of experimental facilities themselves.

Each model type offers distinct advantages for specific HEP applications while sharing the common goal of expanding the capabilities of modern particle physics beyond what traditional analytical and computational methods can achieve. As both machine learning techniques and HEP research continue to advance, we can expect increasingly sophisticated integration of these approaches to address the field's most challenging problems.

References

- [1] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Calogan: Simulating 3d high energy particle showers in multi-layer electromagnetic calorimeters with generative adversarial networks. *Physical Review D*, 97(1):014021, 2018.
- [2] Huilin Qu and Loukas Gouskos. Particlenet: Jet tagging via particle clouds. *Physical Review D*, 101(5):056019, 2020.
- [3] T. R. Harvey and A. Lukas. Particle physics model building with reinforcement learning. *arXiv preprint arXiv:2103.04759*, 2021.
- [4] ATLAS Collaboration. Learning by machines, for machines: Artificial intelligence in the world's largest particle detector, 2021. Accessed: 2025-05-15.
- [5] Geant4 Collaboration. Geant4: A toolkit for the simulation of the passage of particles through matter. <https://geant4.web.cern.ch/>, 2024. Accessed: 2025-05-15.
- [6] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler. Lorentznet: Lorentz equivariant graph neural network for particle physics. *arXiv preprint arXiv:2305.12149*, 2023.
- [7] Konstantin Shmakov et al. Graph neural networks for charged particle tracking at the lhc. *arXiv preprint arXiv:2201.08187*, 2022.