How far have we gone and can we go with machine learning in particle physics?

Timea Vitos

ELFT seminar
September 30, 2025

Financed by
the Swedish Research Council (Vetenskapsrådet),
employed by
Uppsala University
hosted by
ELTE and UCLouvain

Today's talk

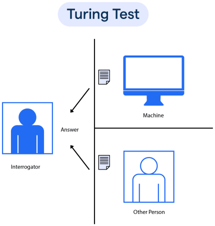
Today's talk

ELTE EÖTVÖS LORÁND UNIVERSITY

We are speaking of machine learning, NOT artificial intelligence!

"Learning: is any process by which a system improves performance from experience." (Herbert Simon)

- After the world wars, researchers began to work towards artificial intelligence, with **Alan Turing** taking the lead
  → **Turing test in 1950** to classify a computer having real intelligence
- Around 20 years later, researchers realized that machine learning could become an own branch and made a clear distinction
- First ideas was to design a computer-algorithm on the human brain (neutrons)
  → **First neural network in 1957**
- In 1997, IBM developed "Deep Blue" - a **winner over the world champion in chess**
- Microsoft, Google, Facebook, Amazon, ...



Caution! Artificial intelligence is a buzzword!

Difference between AI and ML

What is the difference between machine learning and artificial intelligence?

Artificial intelligence (AI) is the broad concept of machines performing tasks that would typically require human intelligence, while machine learning (ML) is a subset of AI that enables systems to learn from data and improve without explicit programming. Think of AI as the overall goal of making machines smart, and ML as a specific technique or toolset used to achieve that goal by learning patterns from data.

Machine learning in a nutshell

Many machine learning models exist!

Decision tree

Boosting

Random forest

Simulated annealing

Logistic regression

Genetic algorithms

Neural networks

Among neural networks, there are also many algorithms:

- Multi-layer perceptron (MLP) **[classification problems]**
- Convolutional neural network (CNN) **[image recognition]**
- Auto-encoders **[missing data imputation]**
- Generative adversarial network (GAN) **[creating new data output]**
- . . .

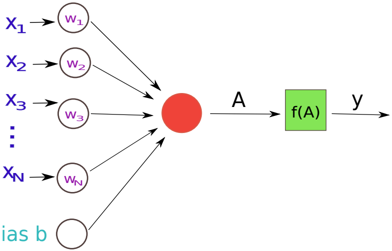
Neural networks in a nutshell

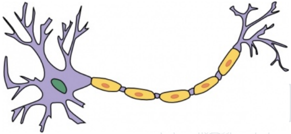
Simplest neural network: the **simple perceptron** (a single neuron cell)

- Input data points x_i
- Synaptic weights w_i
- Add as input a bias b
- Input to the perceptron,

$$A = \sum_i w_i x_i + b$$

- Sum A is passed to an activation function $f(A)$
- Output is y





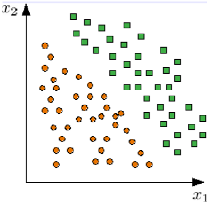
Neural networks in a nutshell

1. Training the perceptron: taking a dataset $\{x_P, d_P\}$ **(training data)**
2. Train the machine by minimizing a loss function, as an example:

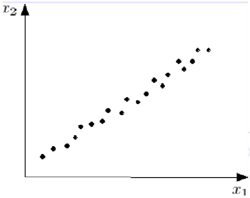
$$E(w_i) = \frac{1}{P}\sum_{k=1}^{P}(y_k - d_k)^2 \quad (1)$$

3. Minimize the function with **gradient descent** methods
4. Evaluate the correctness with **test data**

classification problems



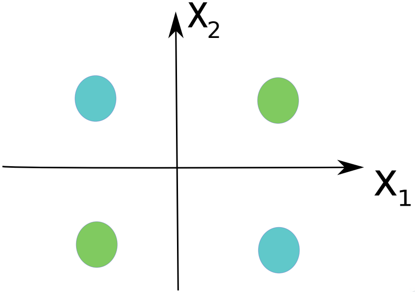
regression problems



Neural networks in a nutshell

Simple perceptron works very well for **linear data**

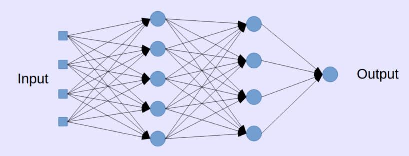
Does not work for any non-linear decision boundaries: XOR problem



Neural networks in a nutshell

Adding hidden layers to network solves the non-linear problem types!

Multi-layer perceptron (MLP)



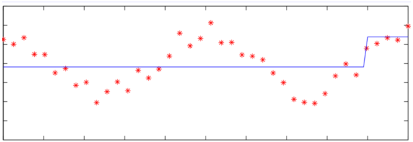
When the layers and nodes in each layers are increased we speak of **deep learning**

(Try out an interfaced MLP for classification problem:
https://playground.tensorflow.org/)

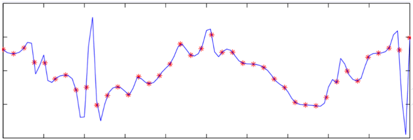
Neural networks in a nutshell

Can we increase the hidden layers and nodes endlessly?

Undertraining (with too few layers)



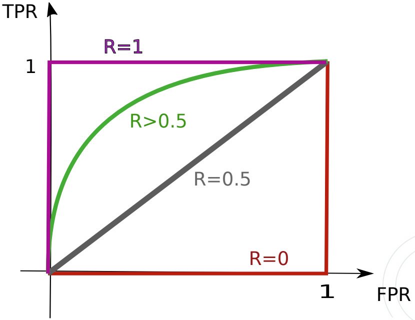
Overtraining (with too many layers)



Performance measure

The ROC AUC score

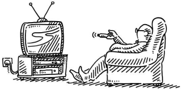
- Plot true positive rate (TPR) versus false positive rate (FPR)
  $\rightarrow$ Receiver Operating Characteristic (ROC) curve
- The area under the curve (AUC) yields a score:
  - Score 0.5: network does not learn anything useful (bad performance)
  - Score $\rightarrow$ 1: good performance
  - Score $\rightarrow$ 0: anti-correlating output



What is the role of machine learning?

Toolkit

- "I have a model and I want to see how well it works!"
- "I do see some pattern but the details are blurry, let's see how a machine tackles this! "
- "My algorithm works, but maybe it can be more efficient with machine learning?"



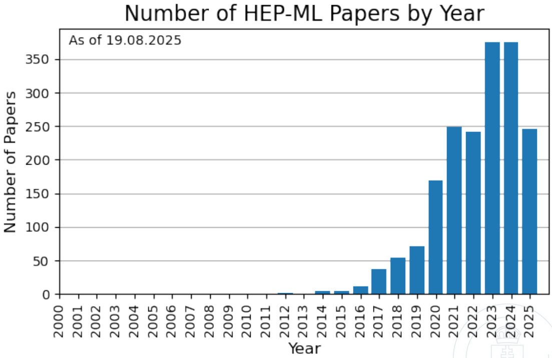
New way of (not) thinking

- "I have a set of data, I could develop models but... the machine can find symmetries faster in it without my effort!"

Today's talk

Current status

For very good updated news and reviews, visit
HEP ML Living Review homepage



Current status

For very good updated news and reviews, visit
HEP ML Living Review homepage

---

[26] arXiv:2509.24092 [pdf, html, other]

Applications of Machine Learning in Constraining Multi-Scalar Models
Darius Jurčiukonis
Comments: 4 pages, 1 figure, to be published in the Proceedings of EuCAIFCon 2025: European AI for Fundamental Physics Conference, Jun 16-20, 2025, Cagliari, Sardinia, Italy
Subjects: **High Energy Physics - Phenomenology (hep-ph)**

[46] arXiv:2509.25169 [pdf, html, other]

Fast, accurate, and precise detector simulation with vision transformers
Luigi Favaro, Andrea Giammanco, Claudius Krause
Comments: To be submitted to SciPost Physics Proceedings (EuCAIFCon 2025)
Subjects: **High Energy Physics - Phenomenology (hep-ph)**; High Energy Physics - Experiment (hep-ex)

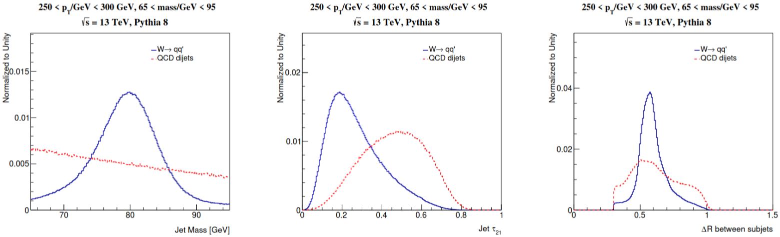
Jet image classifiers

arXiv:1511.05190

Classify jets from boosted particles to differentiate new physics and multi-jet background

---

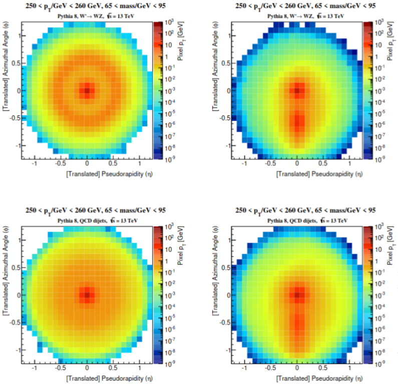
- Discriminate boosted W/Z particles from QCD background mainly with three features: **jet mass**, **n-subjettiness (τ_{21})**, ΔR **between sub-jets**



Jet image classifiers

arXiv:1511.05190

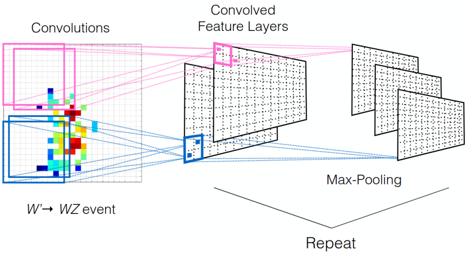
- Investigate the **jet image**: in the (η, ϕ) plane, pixel the average jet intensity
- Generate jet images with simulations (Pythia8) as **test data** for learning
- Consider new physics $W' \to WZ$ process (upper plots) to background (lower)
- Pre-processing of images (left $\to$ right): translation, rotation, re-pixelation



Jet image classifiers

arXiv:1511.05190

- Use a convolution neural network for this **classification** problem
- Optimal kernels are found for (11 × 11)

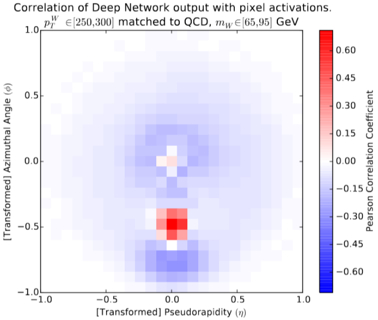


Kernel size	(3 × 3)	(4 × 4)	(5 × 5)	(7 × 7)	(9 × 9)	(11 × 11)	(15 × 15)
AUC	14.770	12.452	11.061	13.308	17.291	20.286	18.140

Jet image classifiers

arXiv:1511.05190

- Network is able to learn the main features of the signal jet image
- Leading discriminant region is the lower jet deposit, and the least correlated is the $(\eta, \phi) \sim (0,0)$ region
- However, this method seems to lose alot of the information in events by doing the projection onto a jet image and requires thus a large amount of data

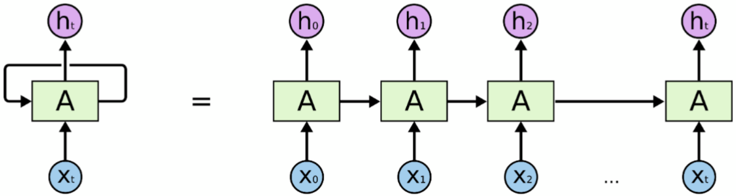


Jet clustering with recurrent networks

arXiv:1702.00748

Can we do better for jet classification than the jet image setup?

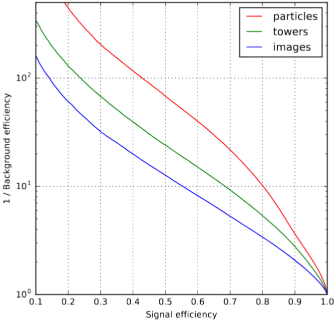
- Jet clustering is a non-deterministic algorithm: the number of jets at the end of the clustering sequence **varies for different random numbers**
- Variable-length input → use **recurrent neural networks** (RNN) used traditionally for language models
- Architecture of the network is built as the sentence is parsed (in language models)
- Recurrent networks can be unfolded to a **feed-forward network** with fixed length

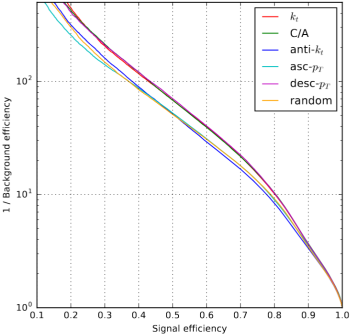


Jet clustering with recurrent networks

arXiv:1702.00748

- RNN classification ("particles" and "towers") outperforms the network based on jet image analysis
- Analyze the (inverse) ROC curve for various clustering algorithms
- The commonly used k_T-algorithm and Cambridge-Aachen are best performers





Today's talk

Kept for a later seminar :)

Today's talk

Today's talk

How far can a machine be learned?

What is the main difference?

Lifeless → follows laws of physics	Life → follows laws of biology
Physics: principle of least action	Biology: principle of largest action
Maximizing entropy	Minimizing entropy
Ideas come from signals	Ideas are sparked (from...?)
Creates something from something	Creates something from nothing

Creating something from nothing

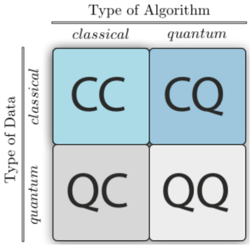
Quantum fluctuations are the closest we can get to in physics
(And not even that can we actually detect...)

- Quantum mechanics opens the platform for many philosophical questions
- Is there more to be learned from quantum data?

- $\rightarrow$ Quantum machine learning!

Quantum machine learning (QML)

- Pioneered in 1990-2000 by **Ventura, Martinez and Trugenberger**
- Combination of machine learning and quantum algorithms can happen in various ways, combination of **algorithm type** and **data type**
- *Personal opinion:* based on the above arguments, QC is the most promising area
- Learn with fewer training data, as the machine can learn several paths simultaneously
- As inherently quantum mechanical systems, **particle physics is an excellent area** to use in this field of research
- This is really cutting-edge topics, so **leave it for a later seminar!**



Today's talk

Summary and outlook

Summary

- Machine learning is widely used in HEP, with exponential increase in the past years
- Used as a toolkit, it can open up new possibilities: in experiments, theory predictions, analysis, phenomenology
- Using it for event generation sub-parts:
  - Improvement is there but not as great as expected for the reweight part of LC-to-FC event generation
  - Improves significantly the importance sampling for integration in MadNIS
- Quantum machine learning is a very new and active field of research, where particle physics and HEP has large potential for being the frontier in machine learning developments

Outlook

- ML4JETS workshop in Vienna, September 14-18 2026, https://indico.global/event/15240/
- . . . and many more will show up!
- IRIS-HEP network (USA), https://iris-hep.org/projects/ml4jets.html
- COST action: Edge Deep Learning for Particle Physics, https://www.cost.eu/actions/CA24153/
- Keep an eye on the Living review: https://iml-wg.github.io/HEPML-LivingReview/

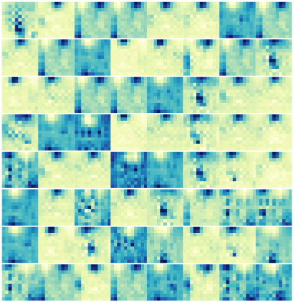
Thank you for listening!

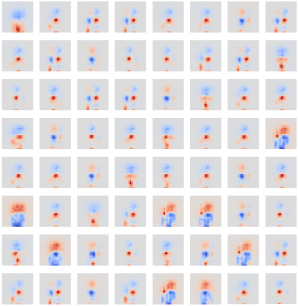
Appendix

Jet image classifiers

arXiv:1511.05190

- Examining the kernels (here only for first layer), one can analyze the **trends** which the machine learns for the various areas in the image
- The machine this ways can reveal what the **main features** it bases its learning on





Jet clustering with recurrent networks

arXiv:1702.00748

Feed-forward network built from the sequential clustering of jets