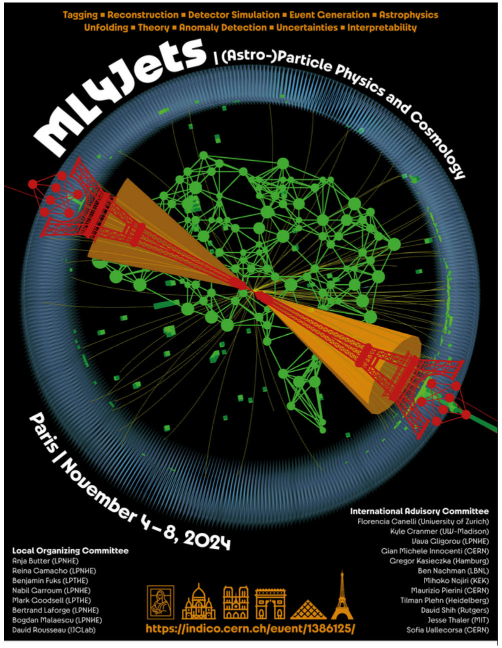
Discussing about

ML4Jets 2024, Paris



ML4Jets

- **ML** conference, not only about hadronic **Jets**
  - It's more ML4HEP
  - also a bit of astro & cosmo topics
- 106 talks, 140 persons
  - plenary & 2 parallel sessions
- I could only attend ~4 days
  - dense but very interesting program
  - vary between report of "production usage" of ML technique in experiment to state of the art R&D

Timetable with
talk slides & videos

Session topics

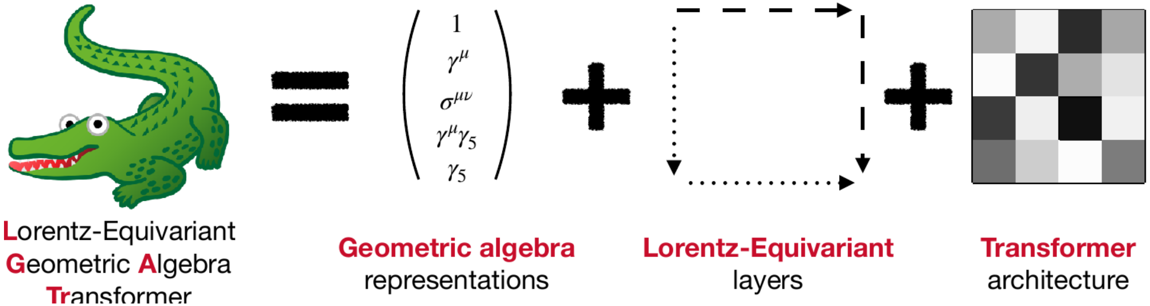
- Event Generation
  - fast parton shower & hadronization with ML
- Simulation & reconstruction
  - Generative NN to speed-up detector simulation
    - Technologies : GAN, VAE, Diffusion models, (normalizing) Flows
  - GNN for tracking at HL-LHC
  - DNN for calibration of electron/γ/jets
- Tagging
  - Identify types/origin of hadronic jets (top vs W/Z vs Higgs vs QCD vs…)
  - Multiple techniques (transformers become dominant)

Session topics

- Astro & cosmo
- Uncertainties & interpretability
  - self/semi/weakly supervised training on dataset
    - ex: "TRANSIT" technique
- Anomaly Detection
- Unfolding
  - various measurements, various tools
    - diffusion/generative models, OmniFold
- Foundation models
  - ex: OmniLearn/Jet, "JetCLR"

Lorentz-equivariant models

- Transformer-based model which is lorentz equivariant
  - i.e Model(boost(x)) = boost(Model(x))
- Lorentz-Gatr model
  - inputs, embeddings are build with elements of a "Geometric algebra" (scalar, vector, pseudo-scalar,... 16 dim)
  - Base architecture very efficient to solve
    - event generation pbm
    - jet classification



Foundation models

Idea :

- pre-train large models on data, unsupervised
  - do it once on very large datasets
- Then fine tune the model for specific tasks
  - can be very quick and/or with small datasets
- Example : OmniLearn based models
  - applications to unfolding (OmniFold), anomaly detction, tagging...

Flow models

- Normalizing flow
  - procedure to learn multidimensional p.d.f with a NN
    - and to be able to sample from them
- Conditional flow matching
  - Fast methods to train & infer "continuous" flow network
  - complex but extremely efficient in generative tasks
    - see talk "The Fast Calorimeter Challenge 2022"
    - also used in unfolding & anomaly detection tasks (?)