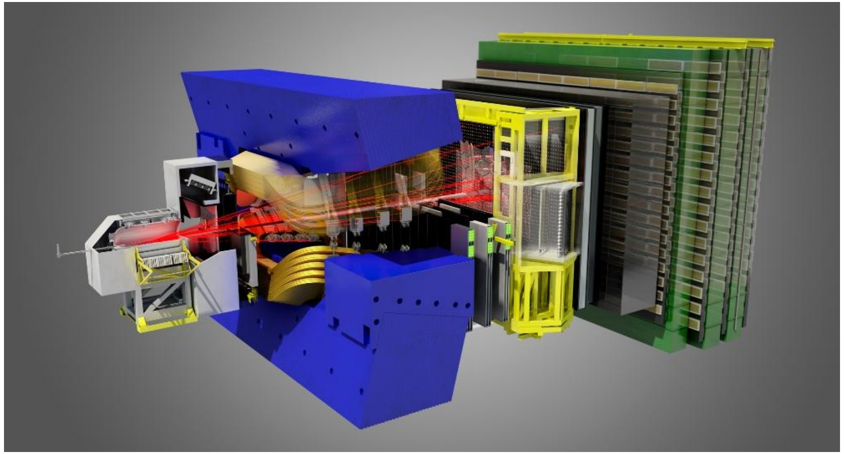
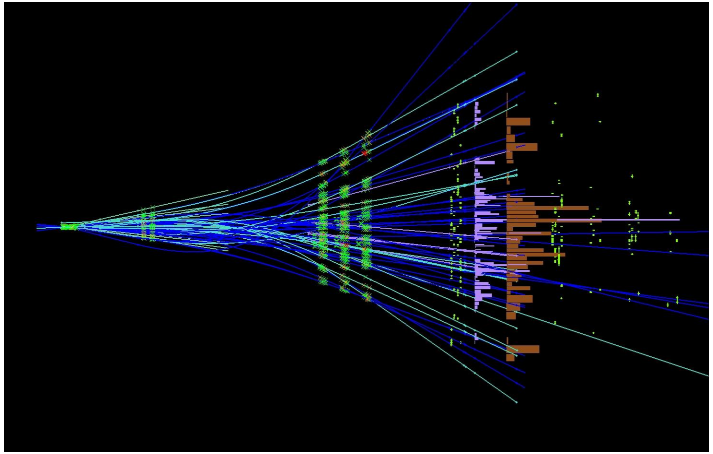
Improvement of GNN-based algorithm for full-event filtering and interpretation at the LHCb trigger

Andrii Koval
Taras Shevchenko National University of Kyiv

Mentor: Jonas Eschle (University of Zurich)

LHCb



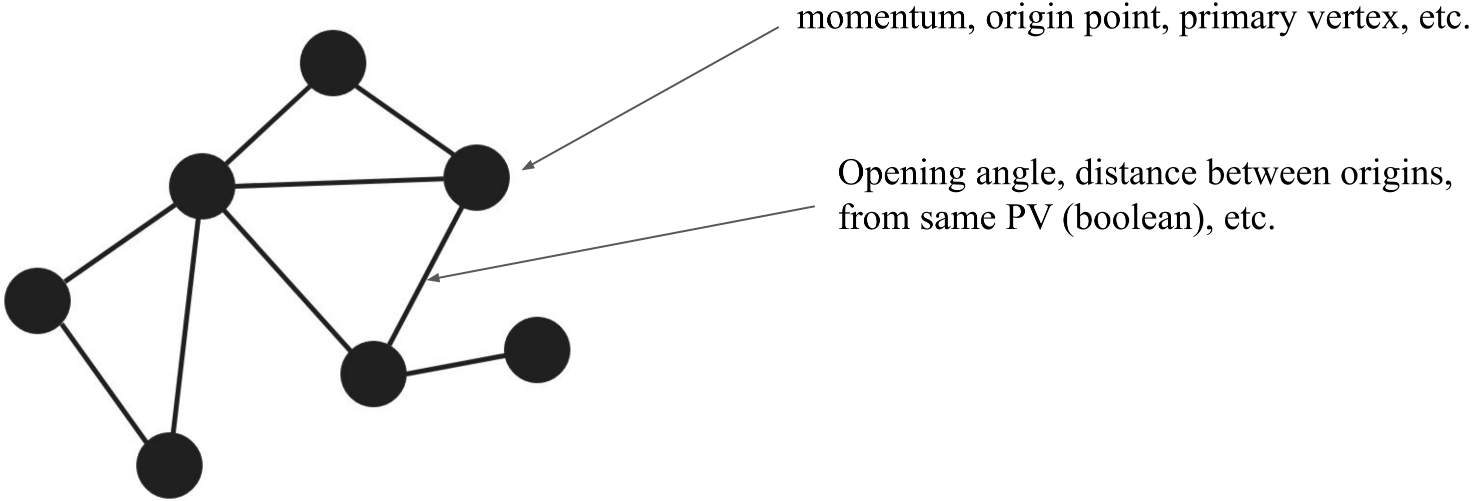


There is a need in reconstructing decay chain

About DFEI group

Create the GNN-based algorithm to:

- Identify particles that come from the decay of a beauty or charm heavy hadron
- Reconstruct the hierarchical decay chain through which they were produced



My role

Goal: Improve time performance of existing algorithm
In practice: Implement and test GravNet* layer on DFEI data

* Based on paper “Learning representations of irregular particle-detector geometry with distance-weighted graph networks” arXiv:1902.07987 [physics.data-an]

Step 1. Learning

1. A Gentle Introduction to Graph Neural Networks (article)
2. CS224W: Machine Learning with Graphs (course)
3. Graph Representation Learning by William L. Hamilton, McGill University (Book)
4. Graph Neural Networks in Particle Physics: Implementations, Innovations, and Challenges (paper)
5. Learning representations of irregular particle-detector geometry with distance-weighted graph networks (paper)
6. Deepmind GraphNets (framework)
7. PyTorch Geometric (framework)

Step 2. Data exploration

Each event (graph) contains ~ 5 proton-proton collisions

~ 6000 graphs

~ 150 nodes (particles) in each graph

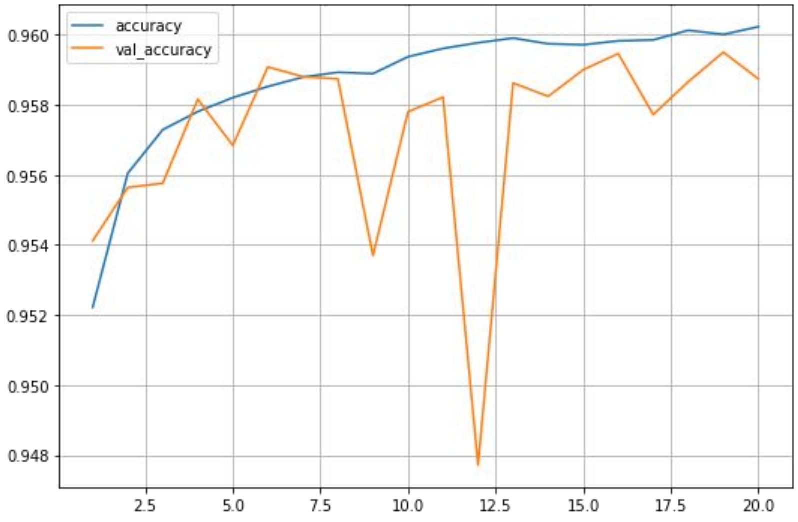
13 features for each node (momentum, origin point, primary vertex, charge etc)

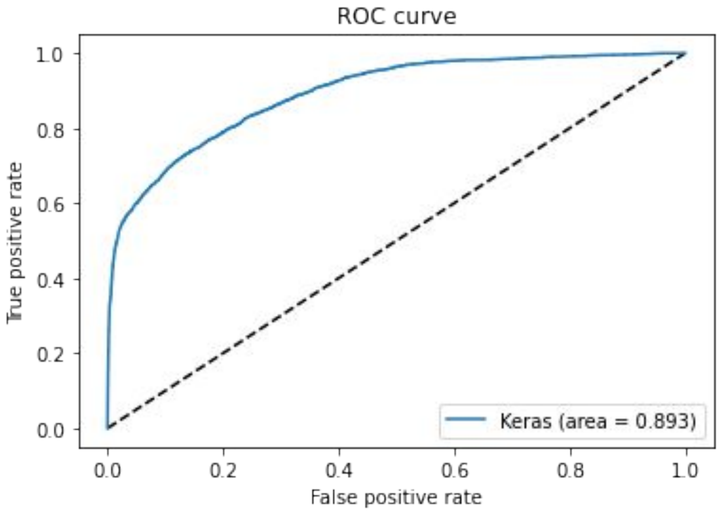
7 features for each edge connecting two nodes (Opening angle, distance between origins, from same PV etc.)

Step 3. First NN on DFEI data

Only node features

5 hidden layers, 300 hidden neurons in each, all use tanh, last – sigmoid





Step 4. Getting familiar with current implementation

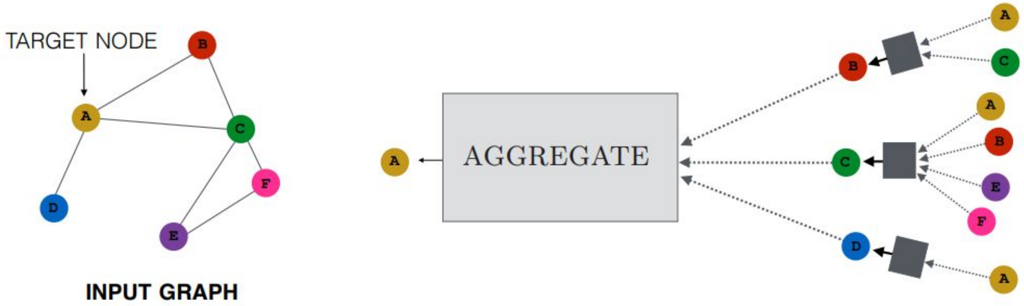
Step 5. Creating dataset

Step 6. GNN first try. Oversmoothing

2 convolution layers

Post MLP (3 layers, ~100 neurons)

Not learning. At all.



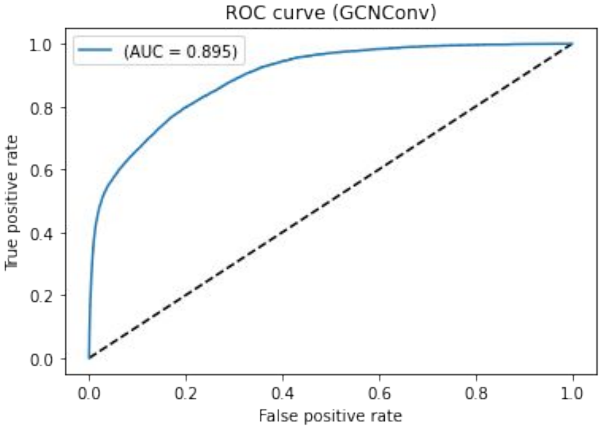
Oversmoothing – a problem that occurs in ML with graphs when nodes “know too much about their neighbours and they all become similar”

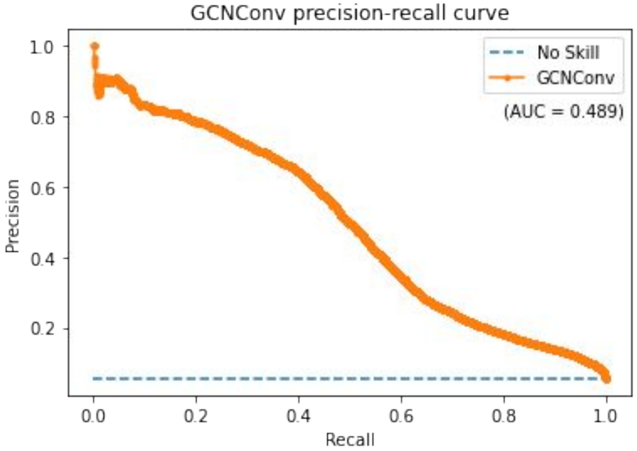
Step 7. Filtering the data

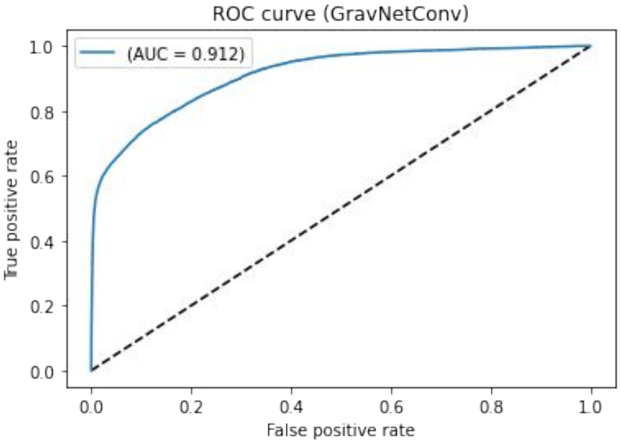
Step 8. Running the model again → Fixing bugs.

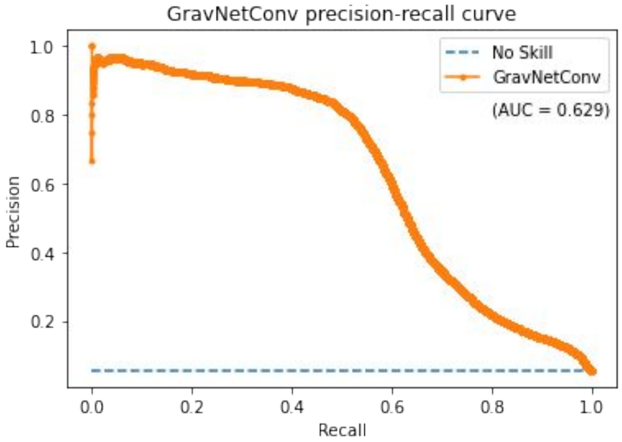
Repeat

Results. Example









Results

	ROC-AUC	Precision-recall-AUC
GCNConv	0.895	0.489
GravNetConv	0.912	0.629
GCNConv 2 layers	0.926	0.597
GravNetConv 2 layers	0.909	0.695
GCNConv (k=20)	0.868	0.432
GravNetConv (k=20)	0.917	0.636

Conclusions

- GravNetConv shows comparable performance to regular convolution at possibly smaller runtime.
- Oversmoothing is an issue, but controllable.
- Bilt code will allow further Hyper Parameters exploration

Big thanks!

- **Jonas Eschle**, my great mentor
- Other members of DFEI group, especially **Julian Garcia Pardinas** and **Andrea Afify**
- **IRIS-HEP**
- **You!**

Git-repository

https://github.com/andrii-hub/IRIS-HEP