Applications of Graph Neural Networks in Particle Physics and Air Quality Systems

Craig Rudolph
Supervisor: Prof. Bruce Mellado-Garcia
Co-Supervisor: Dr. Pallab Basu

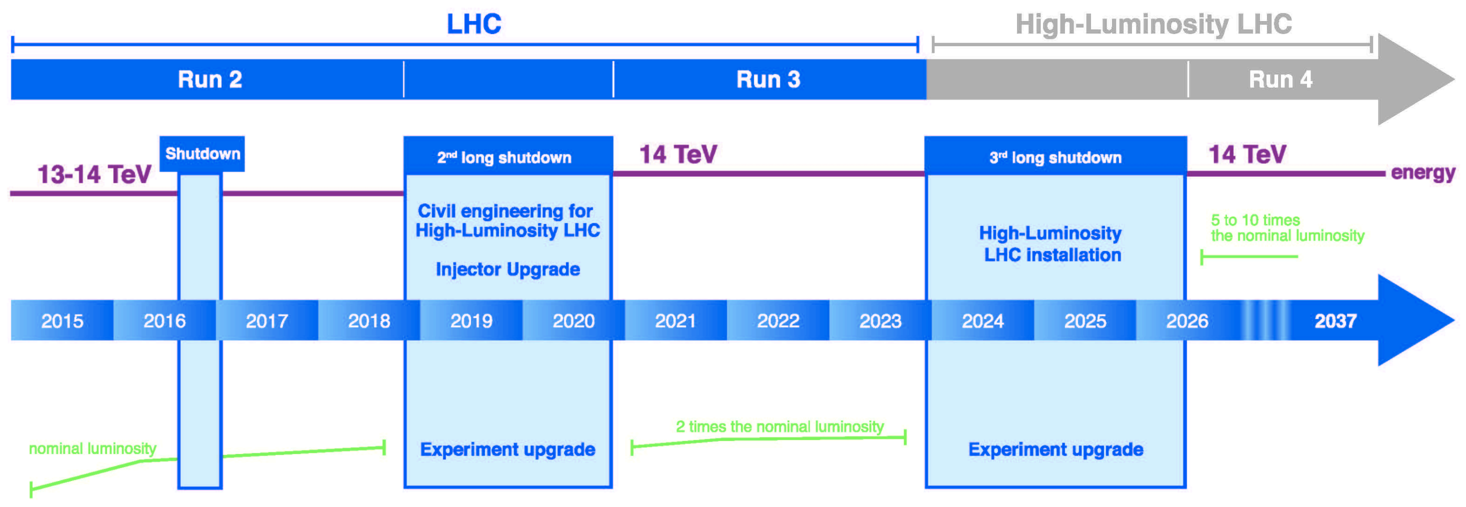
NRF National Research Foundation
iThemba LABS Laboratory for Accelerator Based Sciences

INSTITUTE FOR COLLIDER PARTICLE PHYSICS
UNIVERSITY OF THE WITWATERSRAND

Machine Learning (ML) in High Energy Physics

After the discovery of the Higgs Boson there is an effort to exploit the full physics potential of the LHC and its upgrade HL-LHC which will increase the luminosity by a factor of 5 allowing us to probe the Standard Model and search for new particles. The upgrade will also increase the data density and processing complexity. Therefore, the computing requirements will not be met without significant R&D to ML methods in particle physics to perform these tasks

LHC/ High-Luminosity LHC timeline



Applications of ML algorithms in HEP:

- event and particle identification
- reconstruction
- pile-up suppression

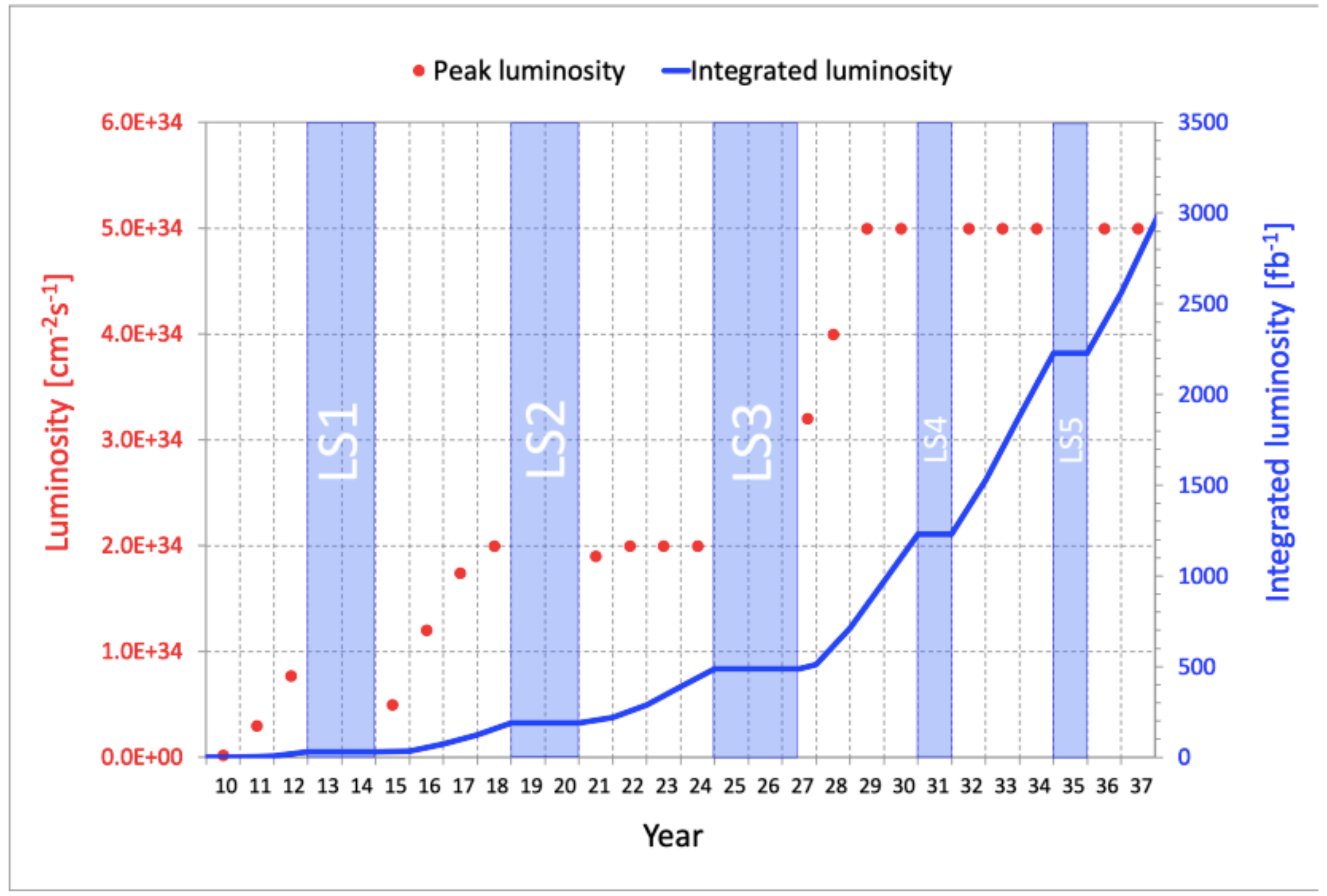
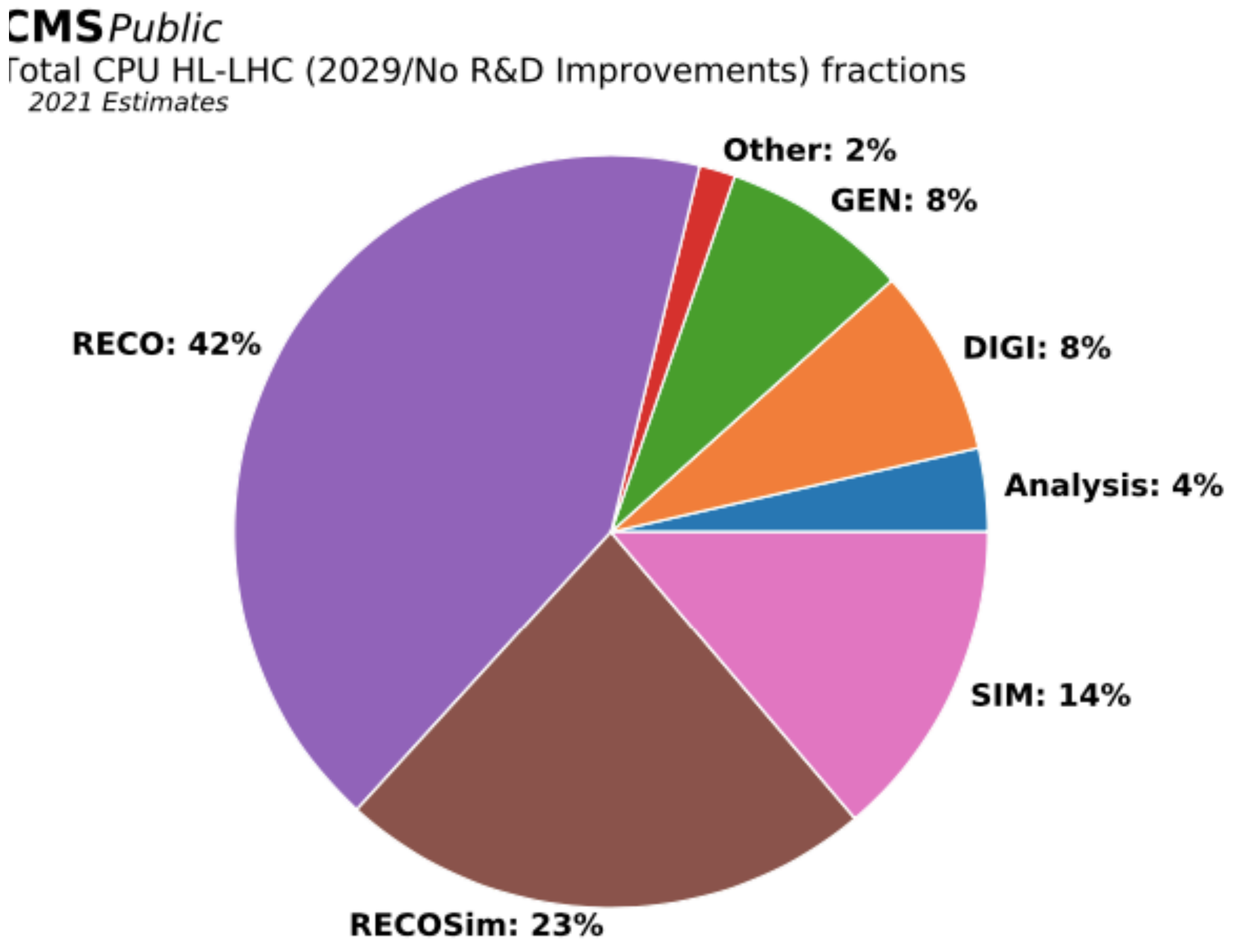
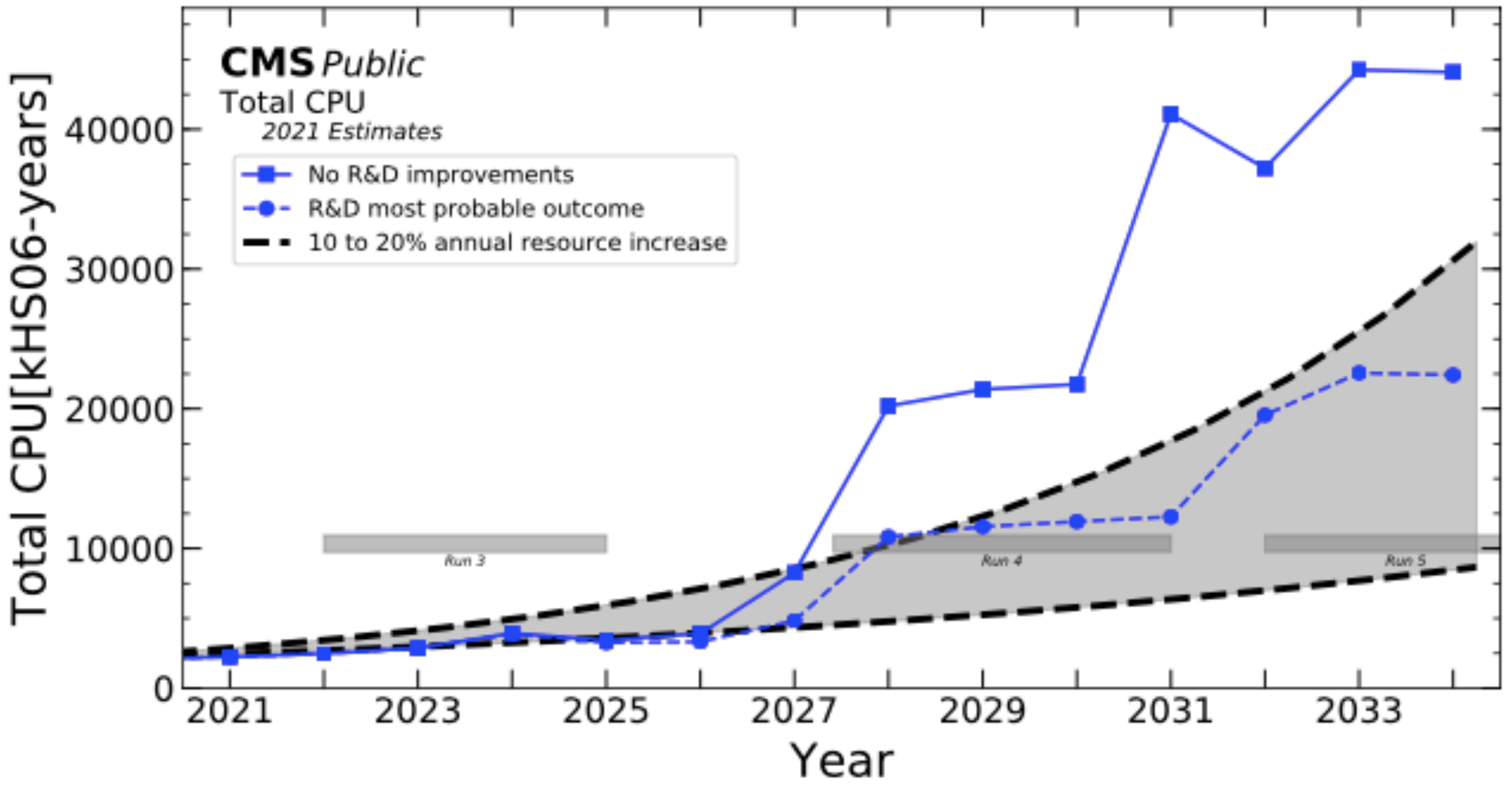


Figure 1. LHC luminosity in Run1 and Run2 compared to predictions for Run3 and for High Lumi





Deep Learning in High Energy Physics

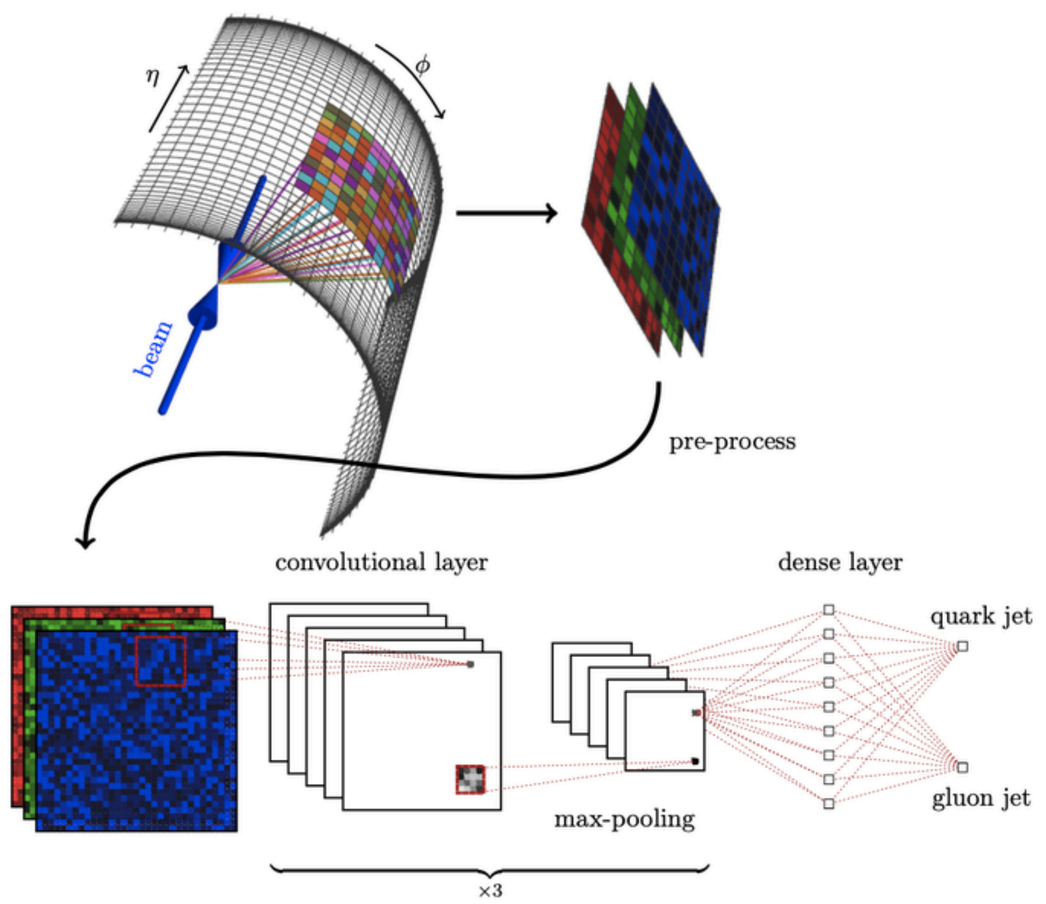
ML Algorithms in HEP: Boosted Decision Trees (BDTs) and Neural Networks (NN)

Deep learning is also used and its full potential is shown when working on low level detector data

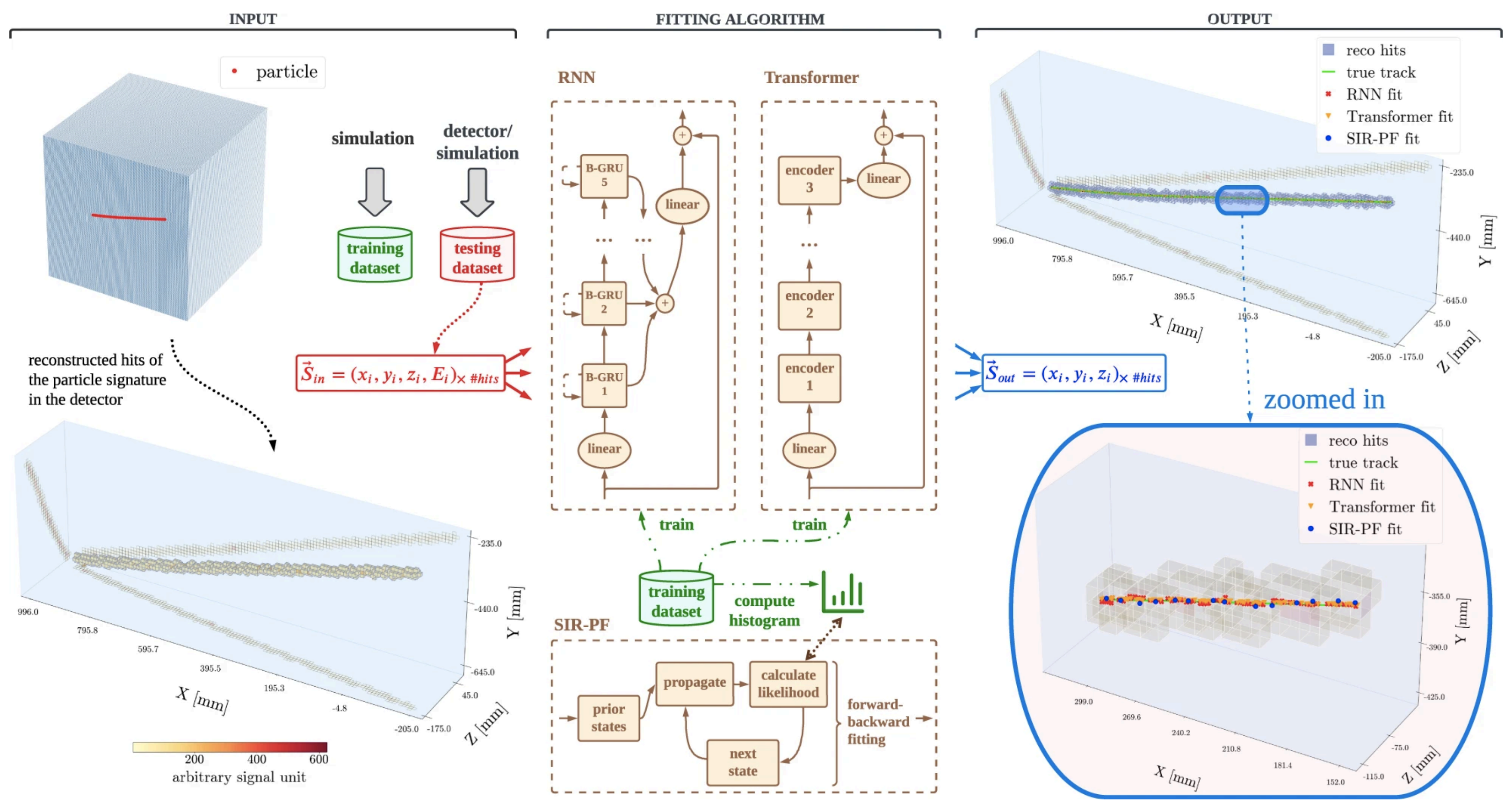
Commonly used DNNs used in HEP:

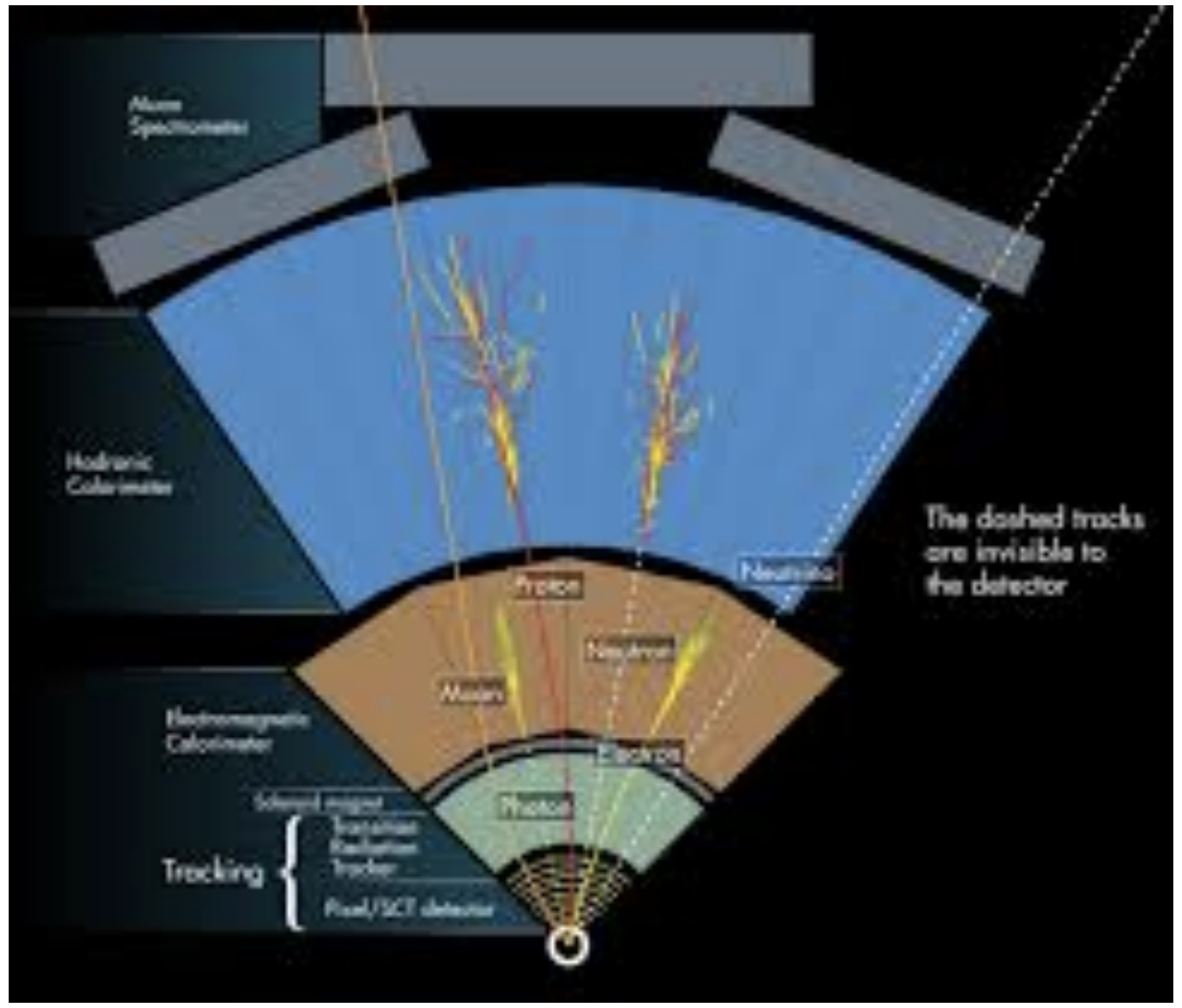
- **convolutional neural networks (CNN) - data is interpreted as images**
- **recurrent neural networks (RNN) - measurements and reconstructed objects can be viewed as sequences**

CNN



RNN





The signals from the detector are heterogeneous in nature due to the different measurements performed by the components of the detector and its irregular geometry. This causes the data from the experiment to be sparse across space and time.

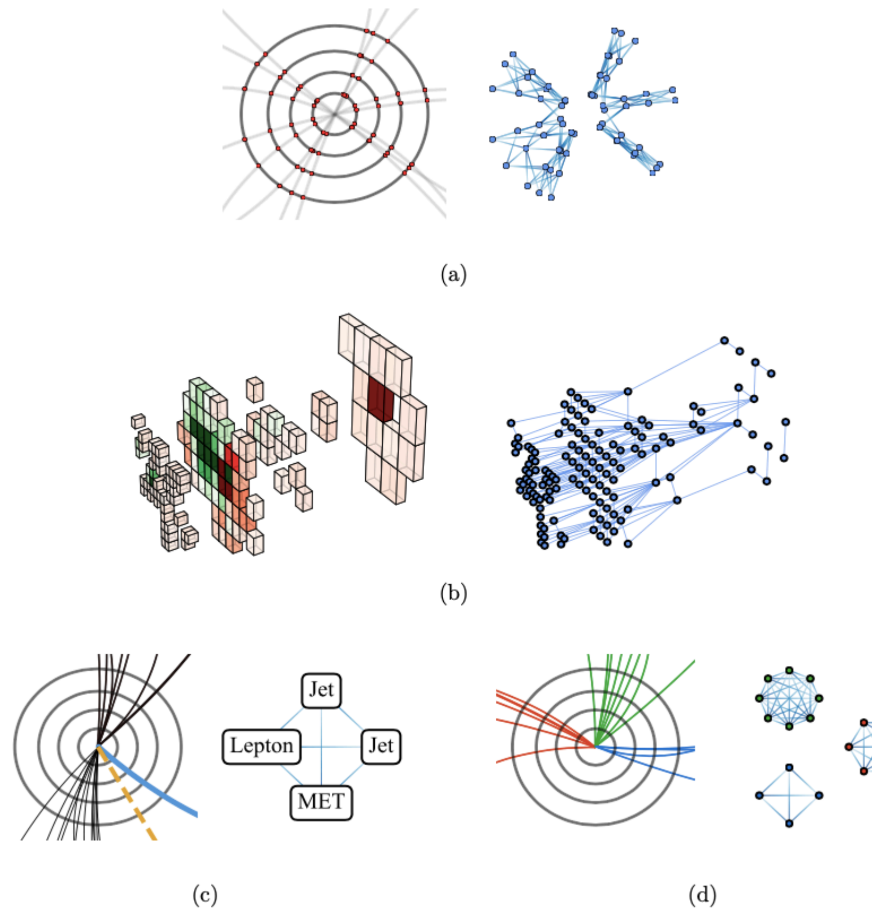
Traditional Deep Learning methods causes information to be lost due to their representation of the data

Graph-based representations and GNN architectures have shown substantial promise for a variety of particle physics tasks

In some cases have demonstrated better scaling properties, reduced resource utilisation, and more efficient data representations compared to traditional methods.

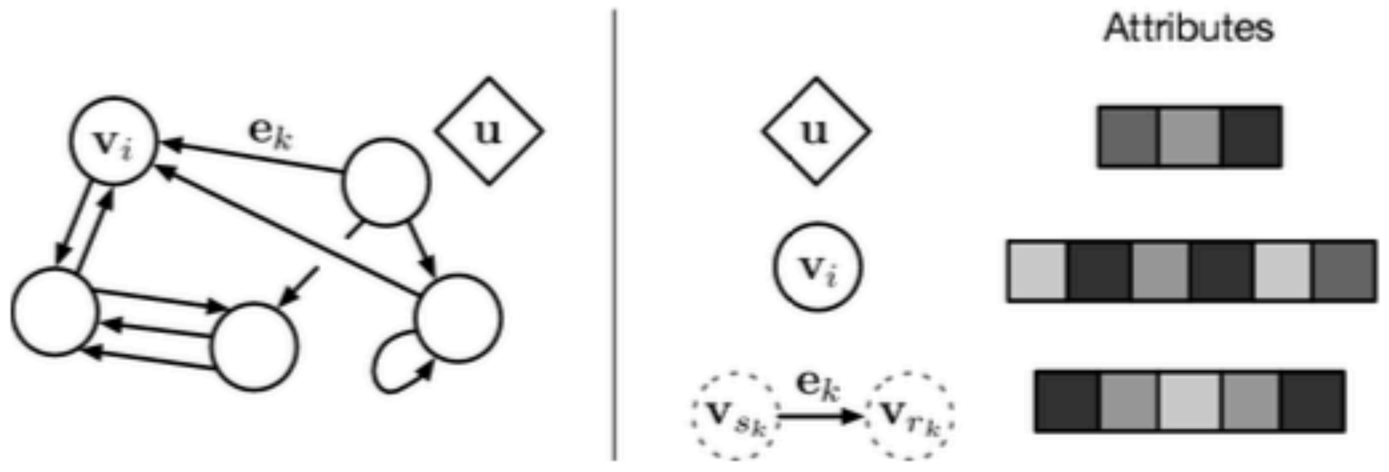
Graph Neural Networks

Data from the detector better represented as unordered set that have relationships which represent the interactions between them which makes the data better represented as a graph.



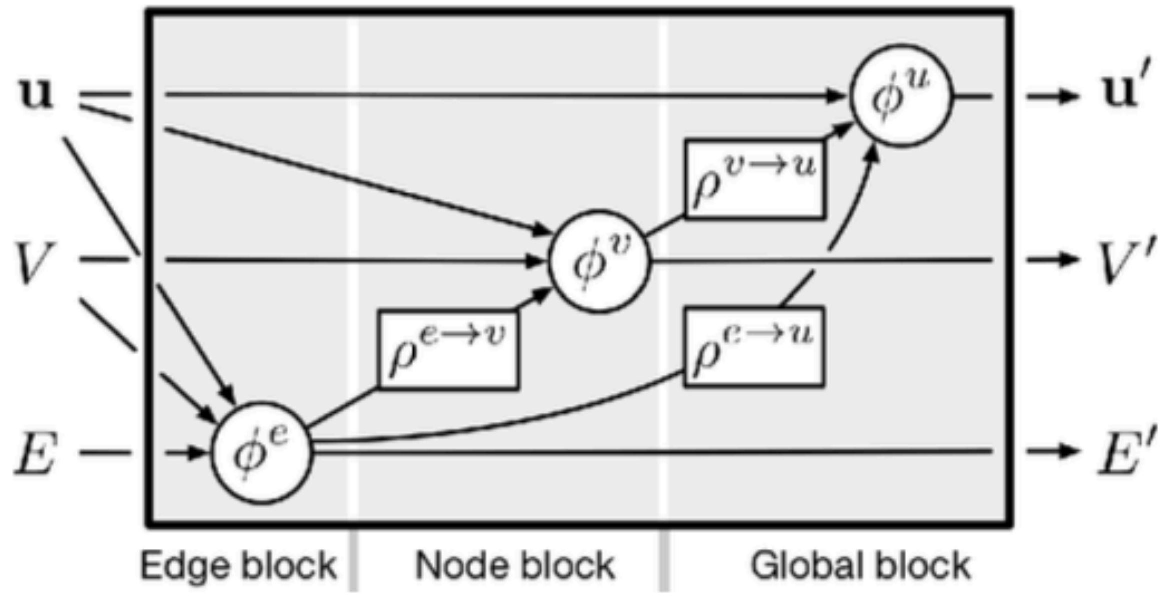
A graph $G = (\mathbf{u}, V, E)$ consists of a set of N^v nodes V, and a set of N^e edges E.

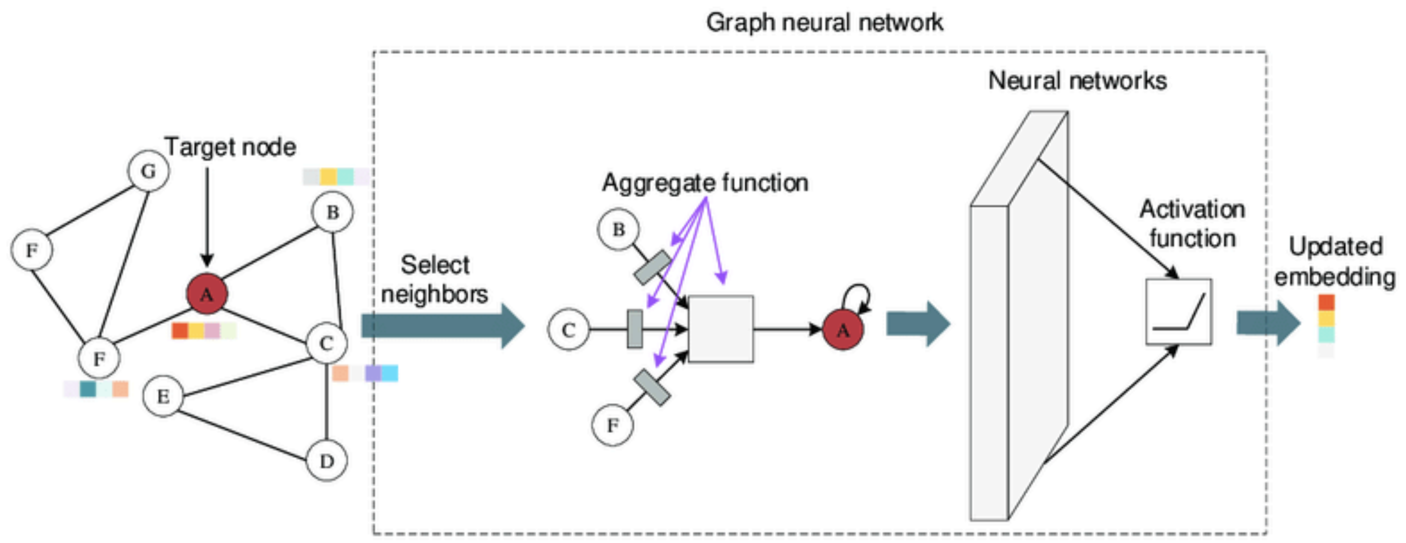
The graph, nodes and edges have their own features. An adjacency matrix defines the edges that connect pairs of nodes



Message Passing

- The information received from the nearest neighbours through the edges in the graph are used to update the node features.
- The node and edge level outputs are each aggregated in order to compute graph-level information in the global block i.e. the graph itself.
- The output of the graph network (GN) is the updated graph that has updated edge, node, and graph features.





The features can be updated either spatially or across spatio-temporal steps where the aggregation and updates share weights with previous time steps.

The Graph Network (GN) blocks are combined to make a Graph Neural Network.

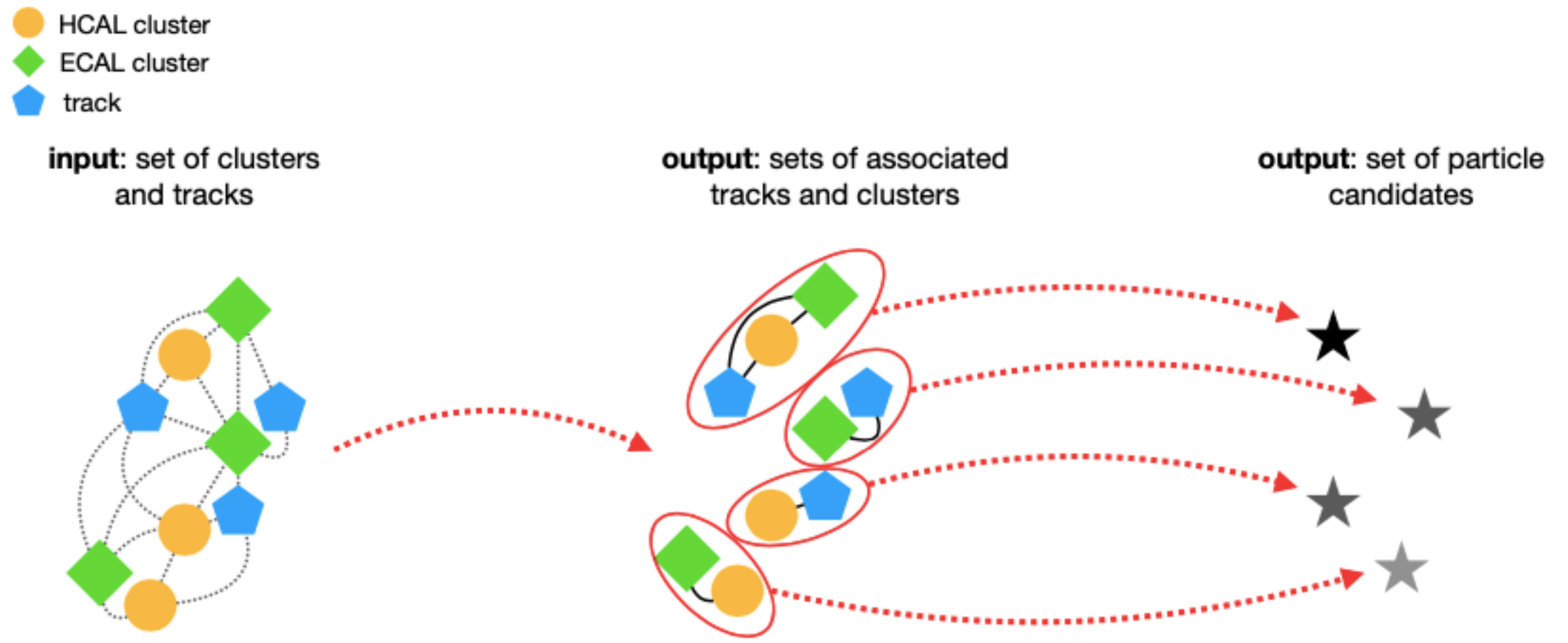
Graph Neural Networks in High Energy Physics

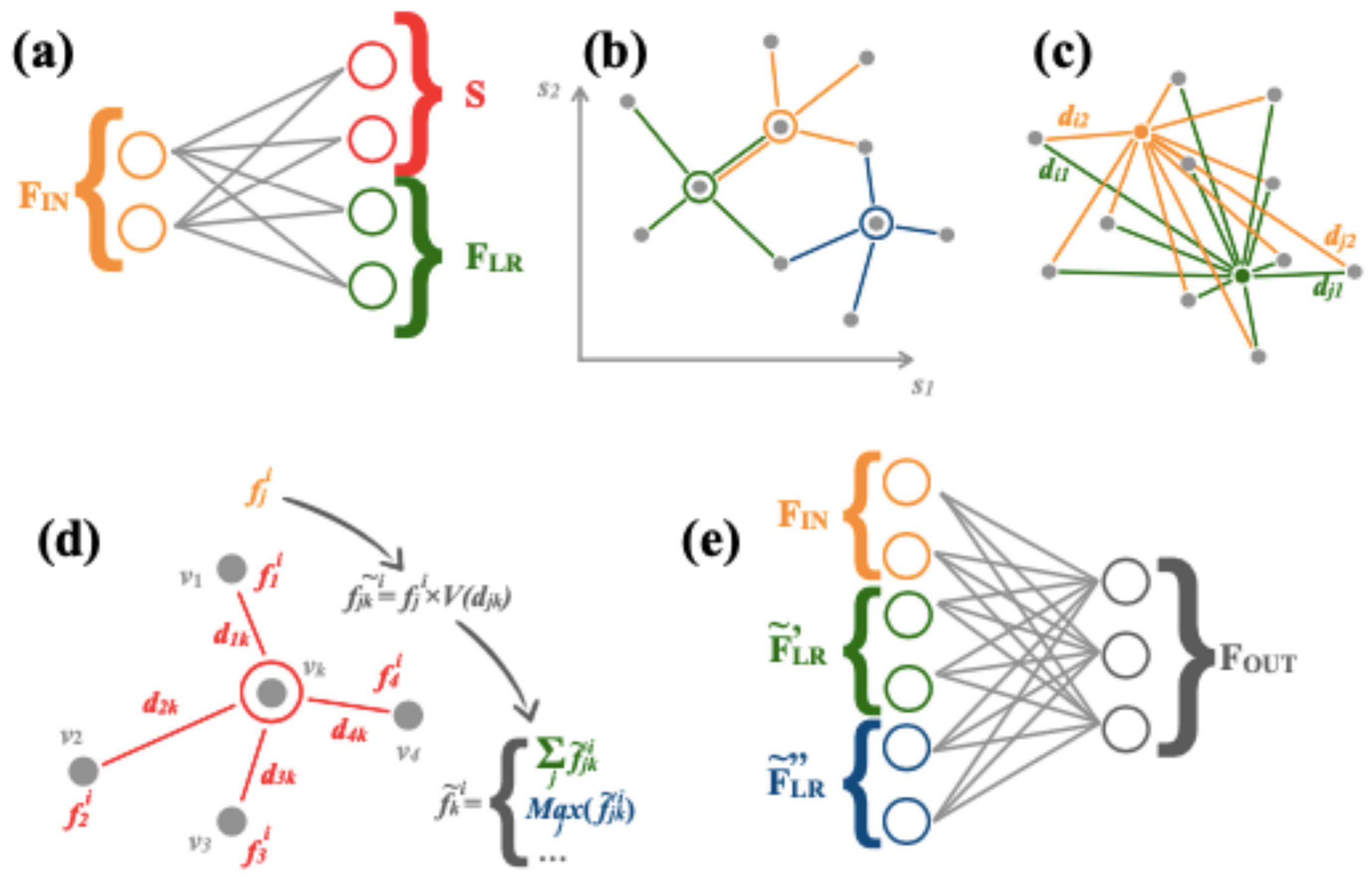
In HEP the majority of the graph-based learning algorithms focus on reconstruction and identification tasks.

For reconstruction the workflow process begins with an unordered set of data points which could be tracker hit spatial points or particle-level kinematic features

The edges are defined either before the GNN or during the learning algorithm. The GNNs produce graph-, edge-, or node-level predictions. In the case of particle flow reconstruction, node-level predictions are made.

Particle Flow reconstruction

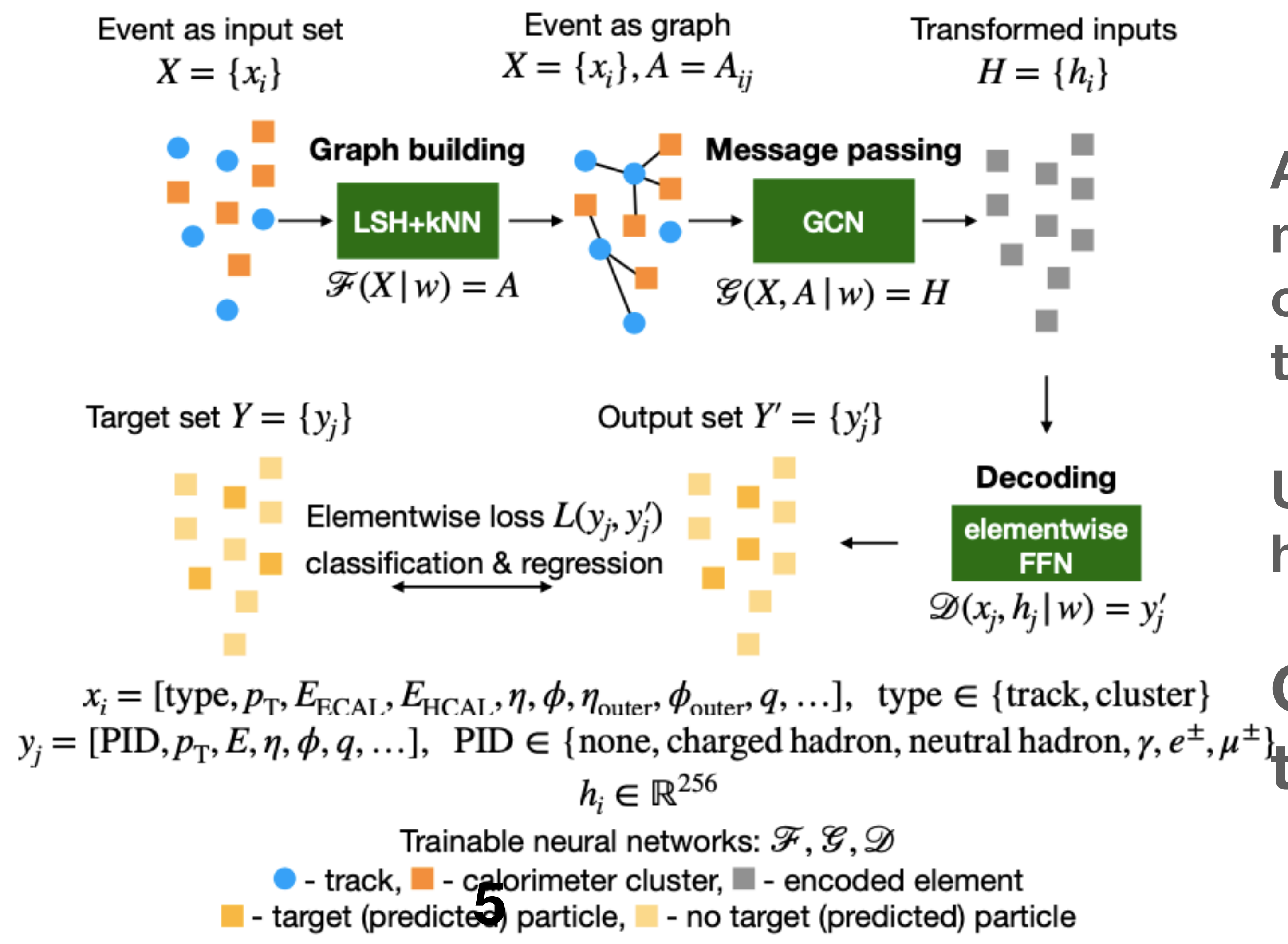




GRAVNET and GARNET

MLPF

GNN used to reconstruct candidate particles in events with a large number of pile up collisions



A set of detector inputs are used for the event and adopt a message passing approach for reconstructing the PF candidates. These reconstructed particles are compared to true values to test the prediction

Underperforms with neutral hadrons and photons in the high transverse momentum region due to pileup

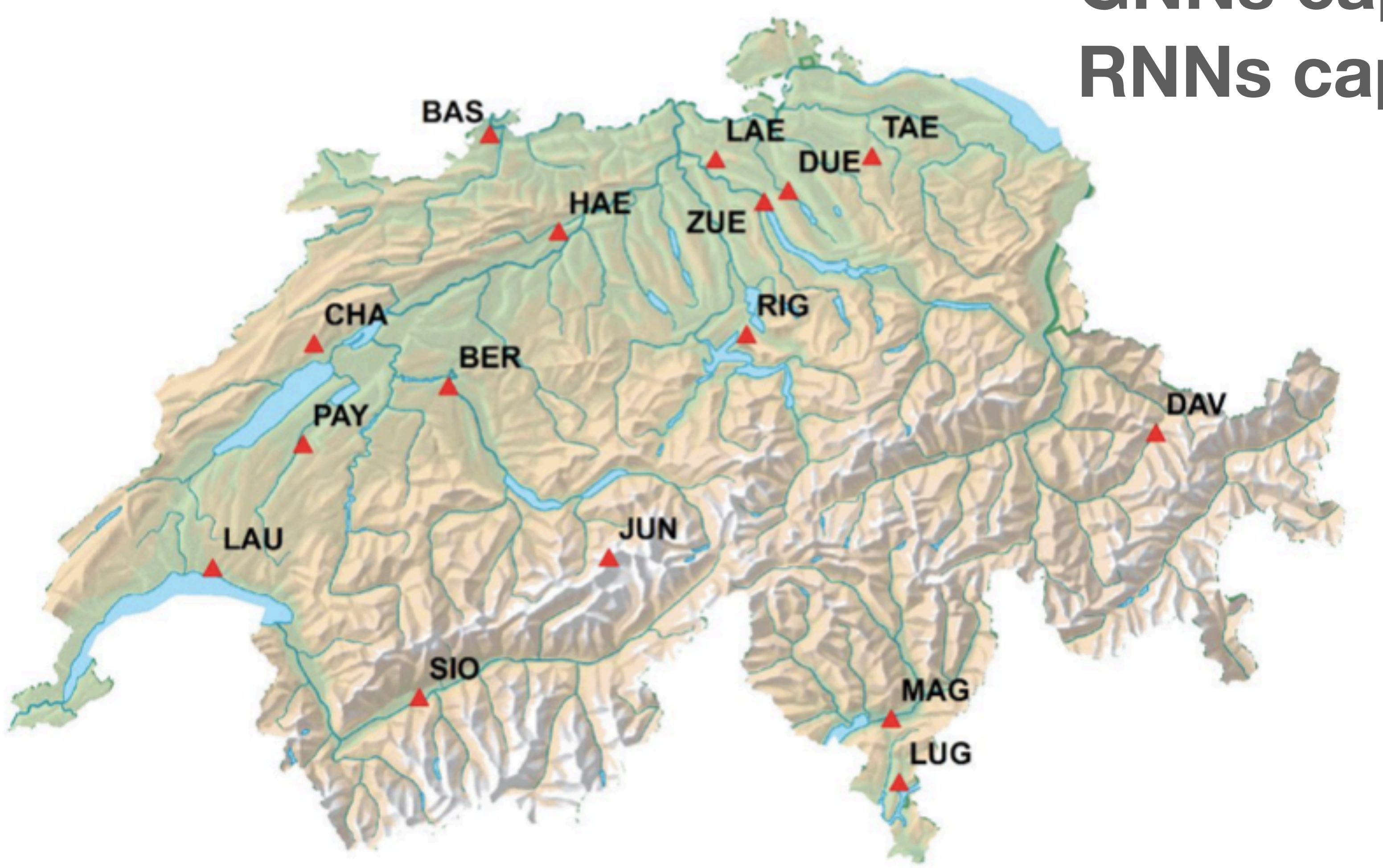
Could be resolved with spatiotemporal GNNs by taking the temporal correlation into account

Spatiotemporal-Graph Neural Networks in Air Quality Systems

Demonstrate the effectiveness of ST-GNNs through the use of air quality systems by combining GNNs with RNNs.

GNNs capture the spatial correlations

RNNs capture the temporal correlations



Station	Code	Longitude	Latitude	Altitude (m)
Basel	BAS	7.58	47.54	316
Bern	BER	7.44	46.95	536
Dubendorf	DUE	8.61	47.40	432
Harkingen	HAE	7.82	47.31	431
Lugano	LUG	8.96	46.01	280
Magadino	MAG	8.93	46.16	203
Payerne	PAY	6.94	46.81	489
Rigi	RIG	8.46	47.07	1,031
Zurich	ZUE	8.53	47.38	409

Graph Construction/representation (G)

Graph - network of air monitoring stations

Nodes - air monitoring stations

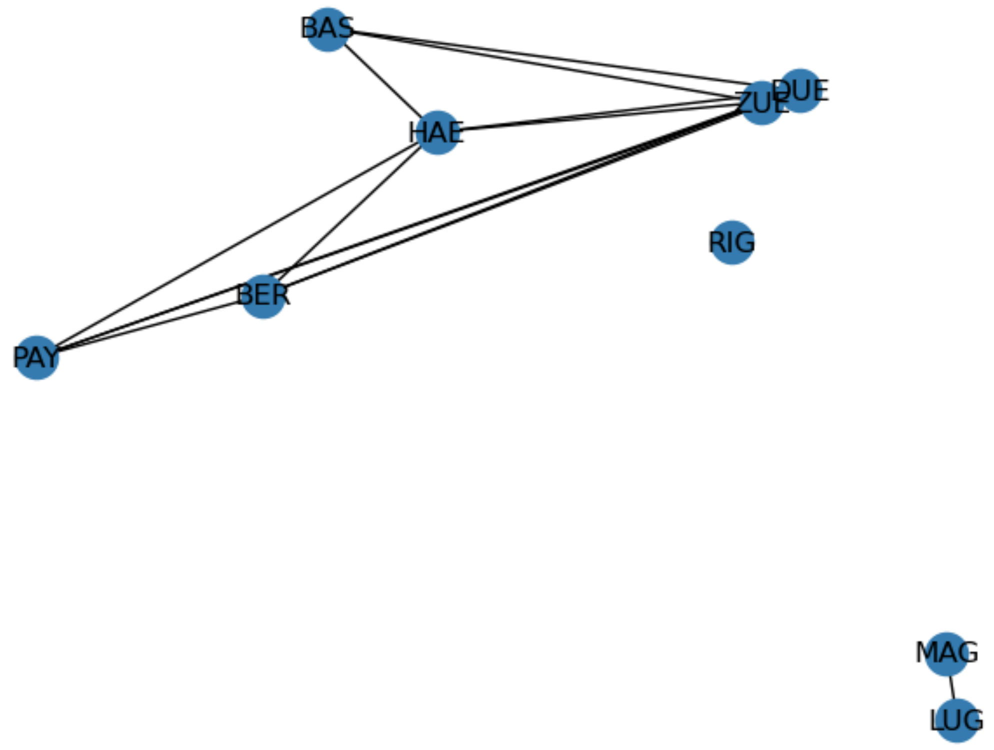
Edges - Interactions between the air monitoring stations

Node Features P

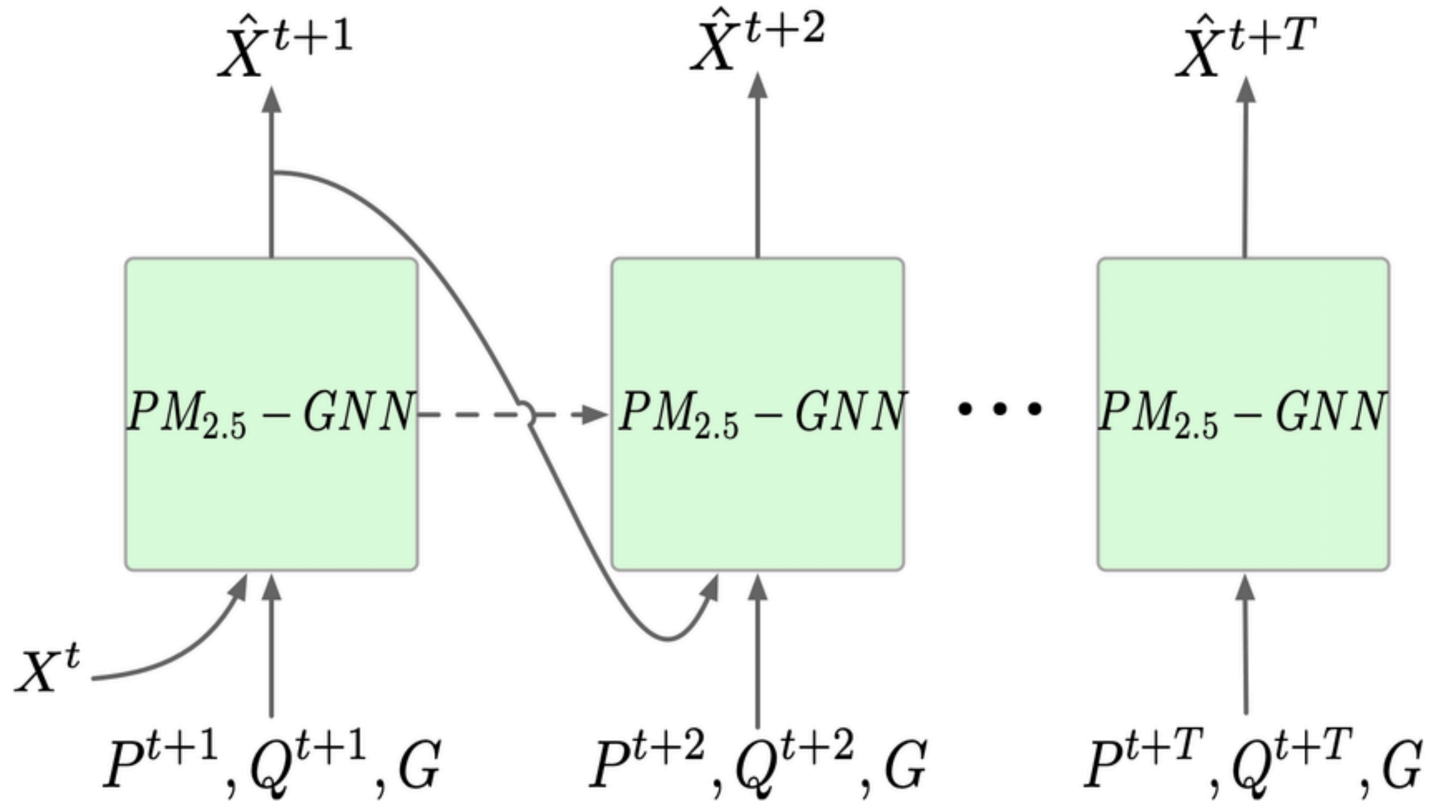
Meteorological features and pollutant concentration data, location coordinates, altitude

Edge Features Q

Non-stationary meteorological features, advection coefficient, distance and direction between other nodes



X - feature we want to predict



Experiment

Datasets

Constructed a 6 year 1 Jan 2016 - 31 Dec 2021 dataset which covers daily measurements made by 9 air monitoring stations

The dataset was divided into a train/validation/test split to examine the model's forecasting ability in a general long-term setting

Train	Validation	Test
1 Jan 2016 - 31 Dec 2019	1 Jan 2020 - 31 Dec 2020	1 Jan 2021 - 31 Dec 2021

Experimental Settings

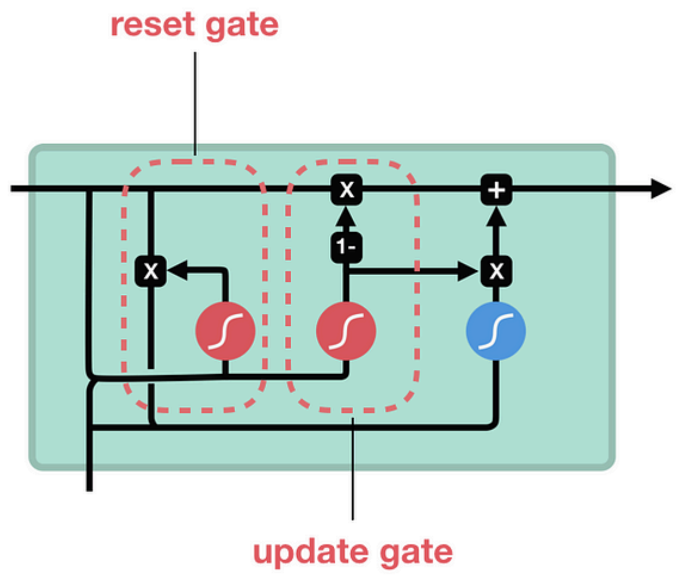
Experiment was ran on 1.8 GHz Dual-Core Intel Core i5
Made use of the PyTorch_Geometric library and its dependencies
All compared models are given identical inputs as well as domain knowledge

Trained using Adam optimiser for 200 epochs
Learning rate and Weight decay are both 0.0005

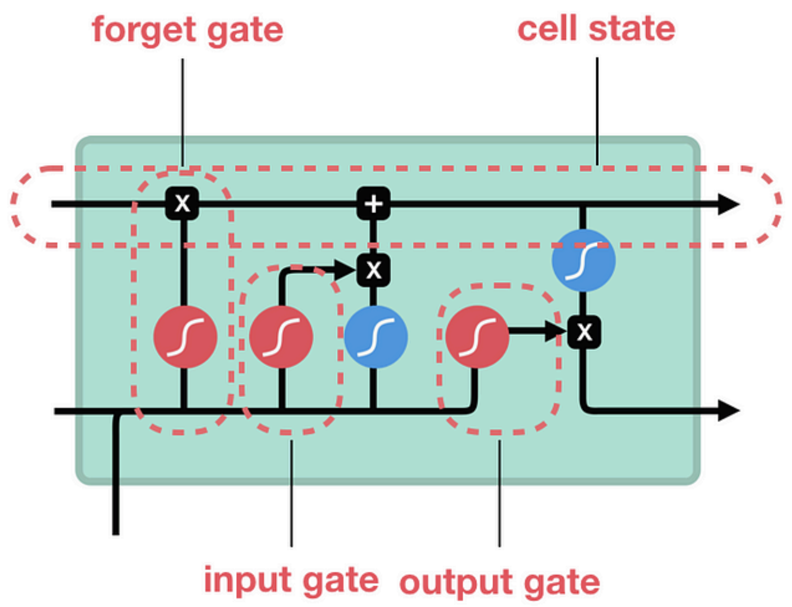
Evaluation of the model performance using the train and test loss to show generalisation ability
Mean absolute error and Root Mean Squared Error to examine the prediction accuracy

Compared Models

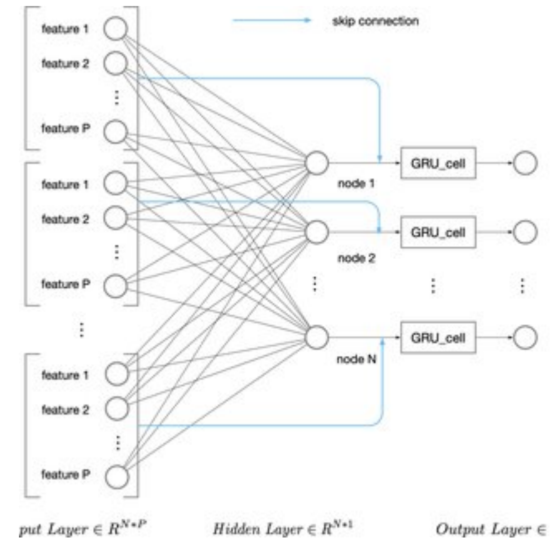
GRU



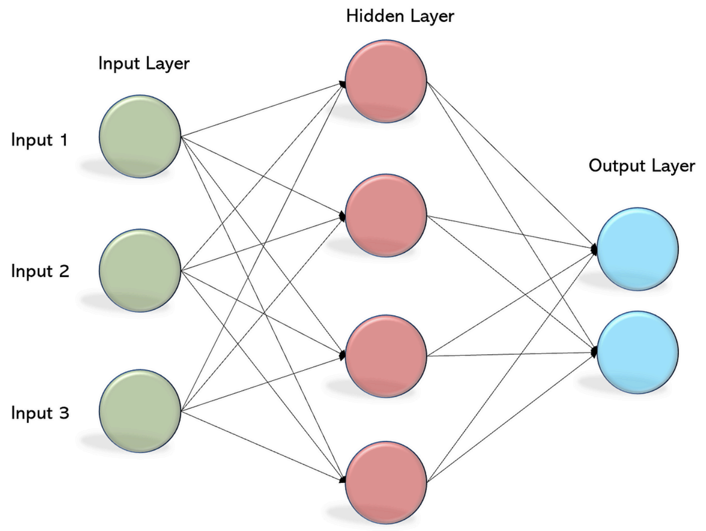
LSTM



nodesFC
-GRU



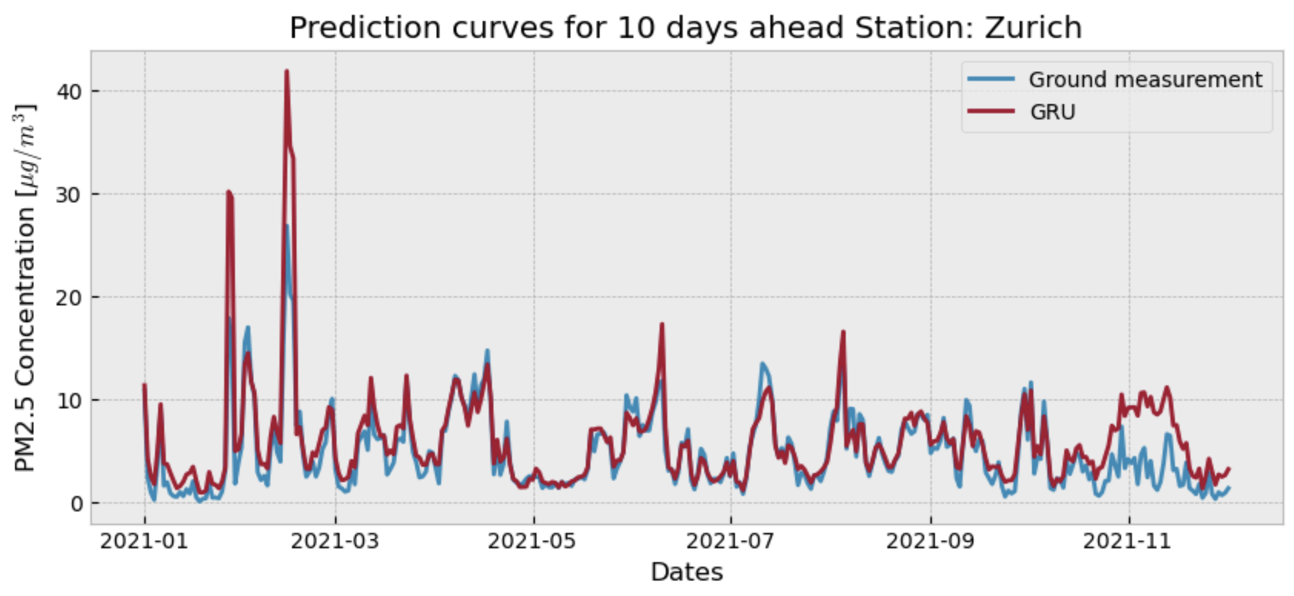
MLP

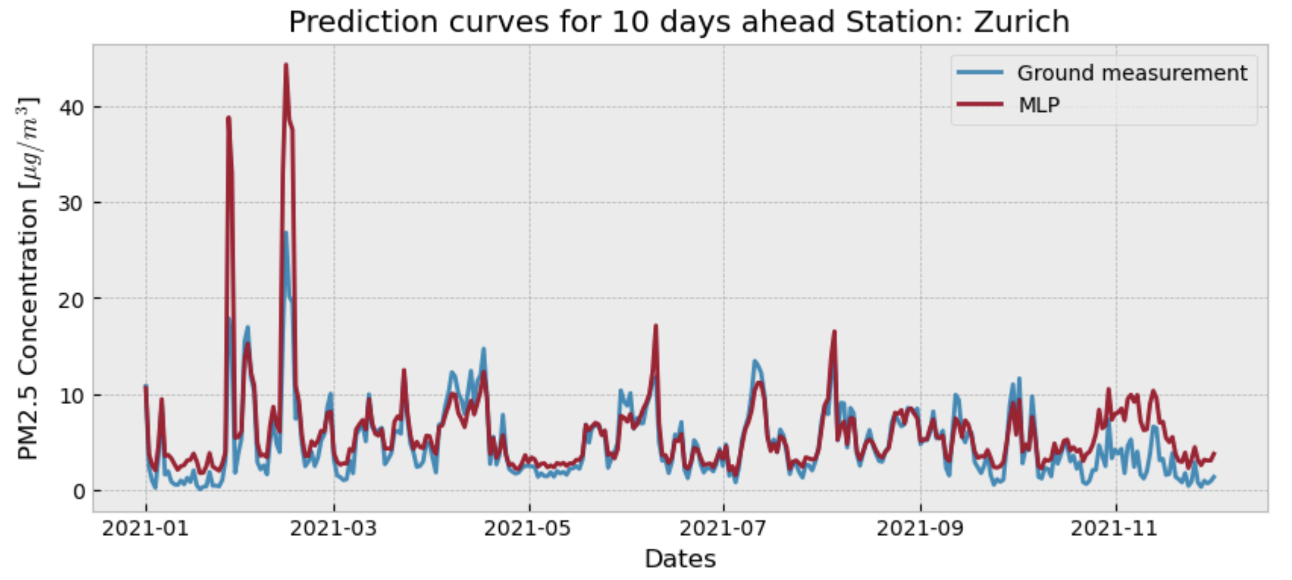


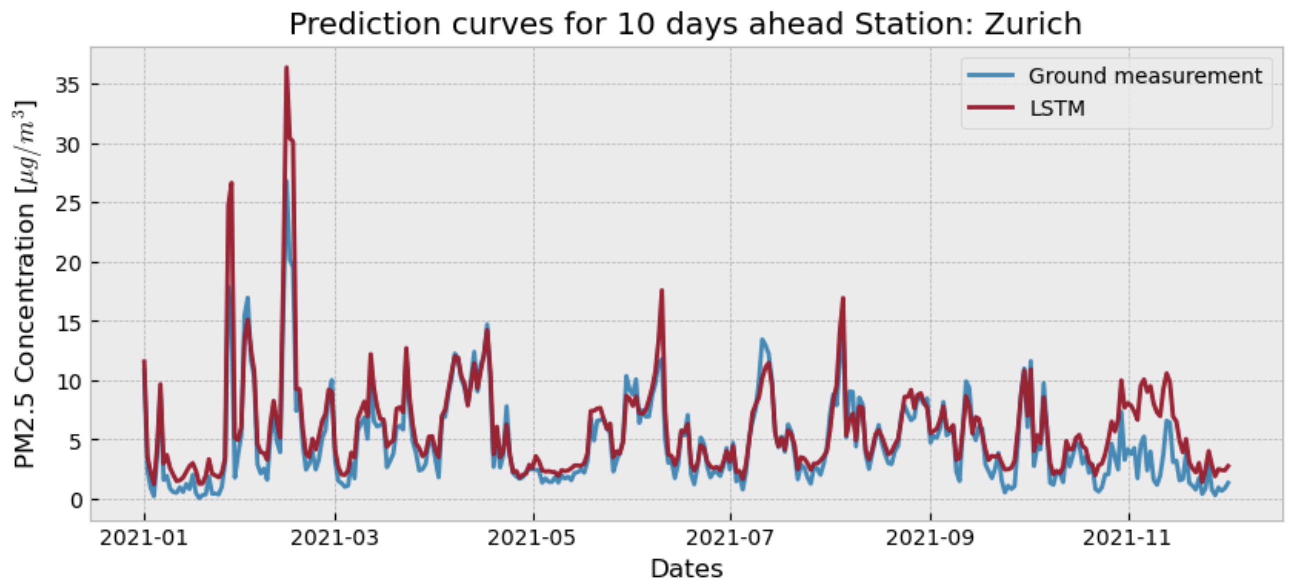
Results

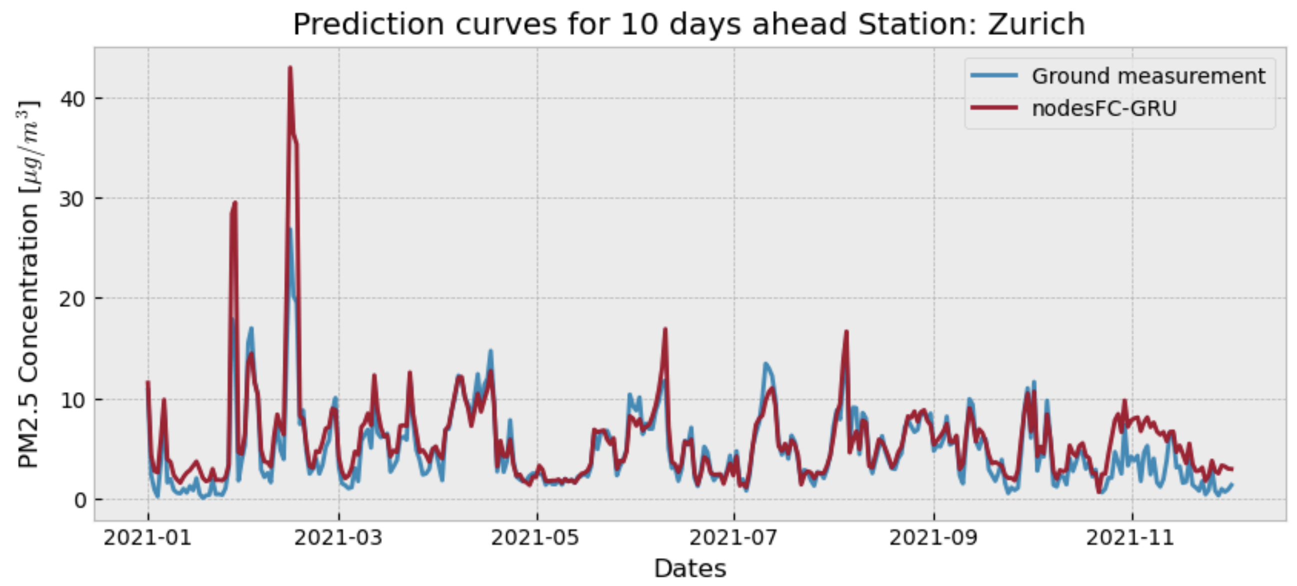
Metric	MLP	LSTM	GRU	NodesFC-GRU	ST-GNN
Train Loss	0.0355	0.0295	0.0283	**0.0212**	0.0236
Validation Loss	0.0445	0.0342	0.0335	**0.0258**	0.0264
Test Loss	0.0571	0.0396	0.0392	0.0384	**0.0301**
RMSE	2.3487	2.0032	1.9976	1.9821	**1.8420**
MAE	1.6709	1.5340	1.5128	1.4558	**1.3338**

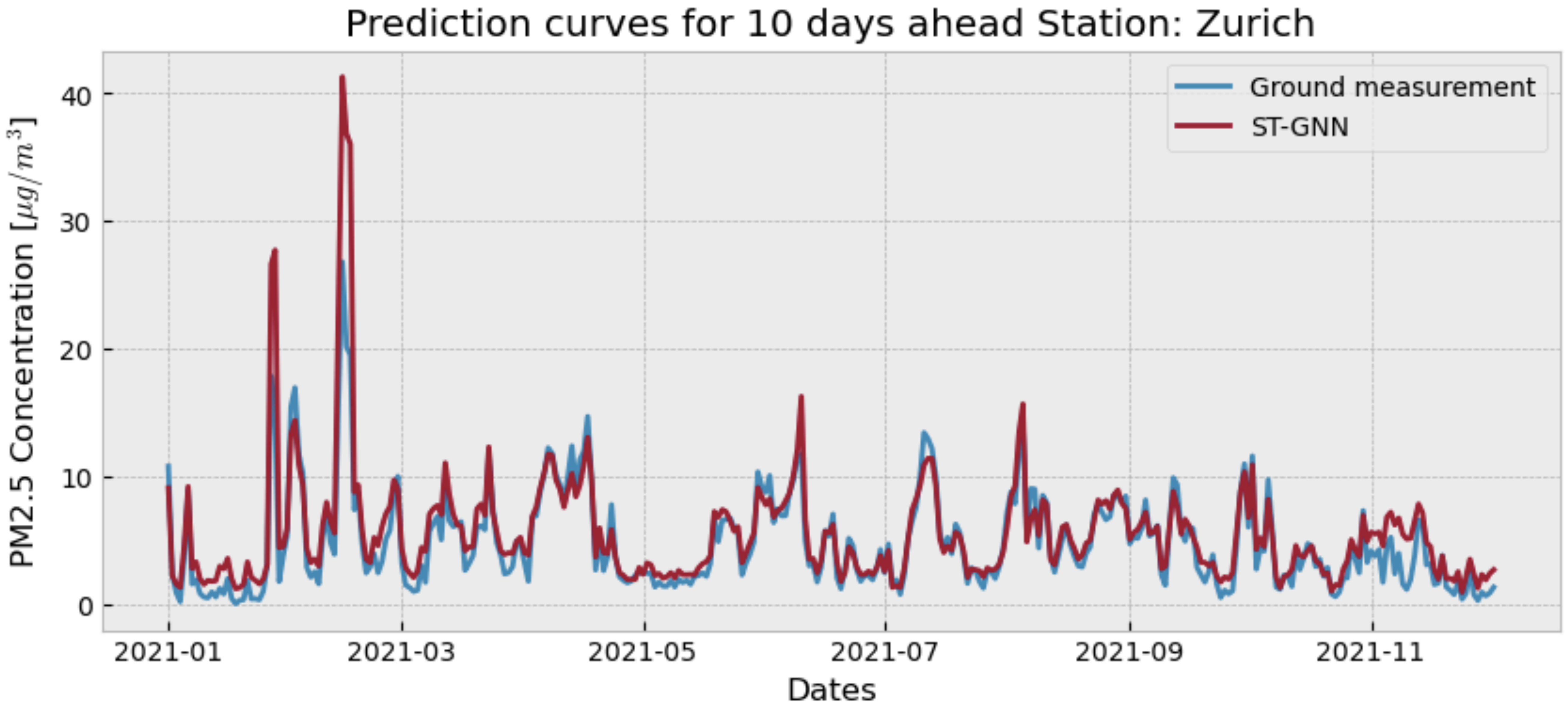
- The nodesFC-GRU and ST-GNN outperform the MLP, LSTM, GRU models.
- NodesFC-GRU suffers from overfitting
- ST-GNN is better at generalising - shown by how close the different loss values are
- ST-GNN outperforms all the other models with its prediction ability shown by the RMSE and MAE











Conclusion and Future Work

Demonstrated the effectiveness of spatiotemporal graph neural networks in air quality systems. This was done to show that it could be used in particle physics tasks in predictive tasks such as physics processes and event reconstruction in which the time dimension plays crucial role.

The next step is to apply this model on ATLAS Run 3 Datasets

There will be ongoing improvement to this model so that it would be better suited for ATLAS data