ML4Pions:
Mixture Density Network Energy Regression

Adithya Suresh, Alex Chkodrov, Saransh Kacharia

Outline

- Purpose of Project
  - Pion Energy Regression
  - Goals of MDN Pion Energy Regression
- Calorimeters
  - EM Calorimeter
  - Hadronic Calorimeter
- Models
  - MDN vs DNN models
- Results
- Conclusion

Pion Energy Regression

Calibrating the pion energy response is a core component of reconstruction in the ATLAS calorimeter.

Machine learning is used to calibrate the pion energy response; data from a GEANT4 simulation of a pion shower in the ATLAS detector is used to train a regression model to predict the pion's energy from the measurements made by the calorimeters.

Goals

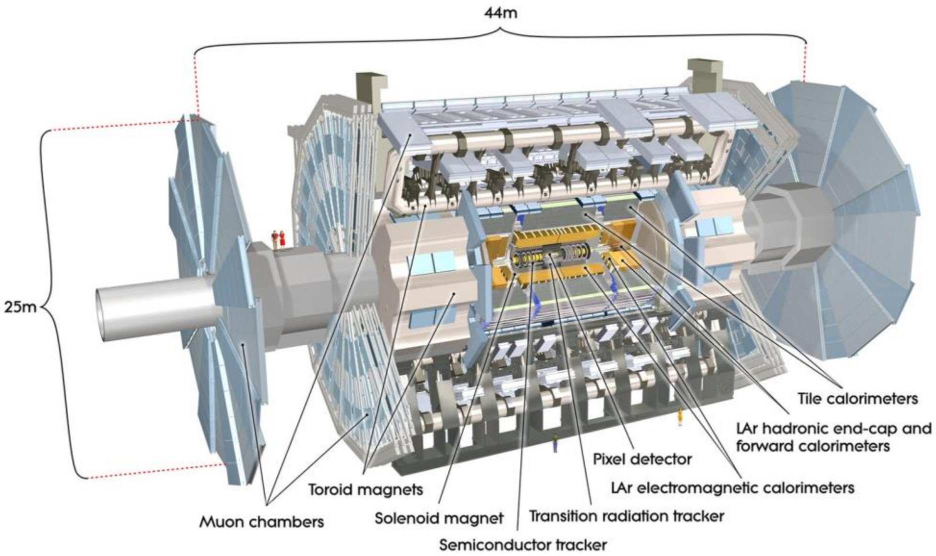
- ★ Recreate PubNote results with Deep Neural Network (DNN)
- ★ Append a Tensorflow Mixture Density layer to the network (MDN)
- ★ Compare MDN (1 component) predicted mean energy to DNN predicted energy
- ★ Possibility of mixing more components in MDN layer to outperform DNN model

Reasons to Test MDN Model

★ Unsure about distribution of energies, so using an MDN model would allow different distributions to try and represent out the data

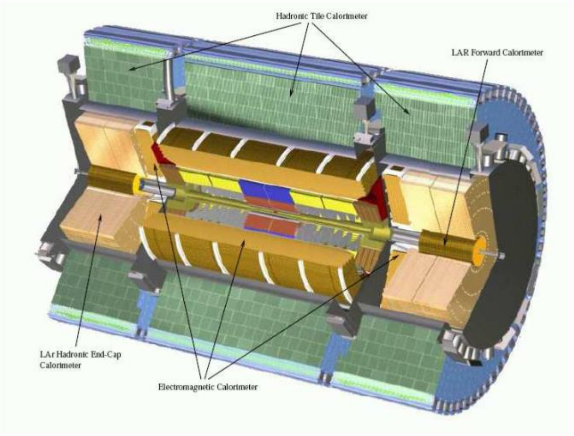
★ MDN's are used to address the underlying uncertainty of a prediction

ATLAS Detector



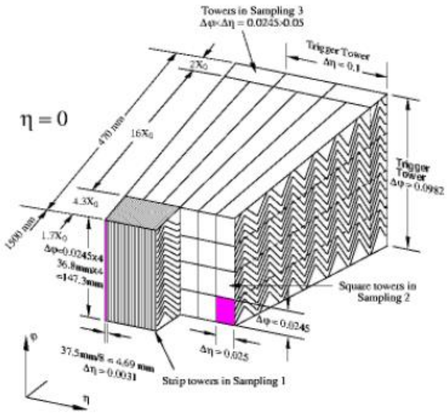
Calorimeters

- ➔ Calorimeters are an experimental device that measures the energy of particles.
- ➔ Particles enter the calorimeter and initiate a particle shower; then the particles' energy is deposited, collected, and measured in the calorimeter.
- ➔ Calorimeters have some amount of uncertainty
  - ◆ ATLAS calorimeters are sampling calorimeters making them less precise compared to homogenous calorimeters



Electromagnetic (EM) Calorimeter

- ➔ Electromagnetic calorimeters are used to measure the energy of electrons and photons as they interact
- ➔ The ATLAS electromagnetic calorimeter is a sampling calorimeter with accordion geometry absorbers immersed in Liquid Argon
    - ◆ EM sampling calorimeter consist of alternating layers of an absorber, a dense material used to degrade the energy of the incident particle, and an active medium that provides the detectable signal
- ➔ ATLAS EM Calorimeter contains a pre-sampler to correct for energy loss of particles in inactive material and three samplers



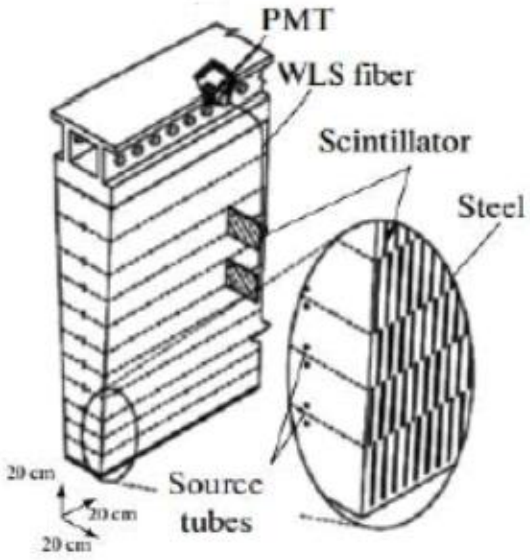
Hadronic Calorimeter

➔ Hadronic calorimeters are used to measure mainly hadrons through their strong and electromagnetic interactions

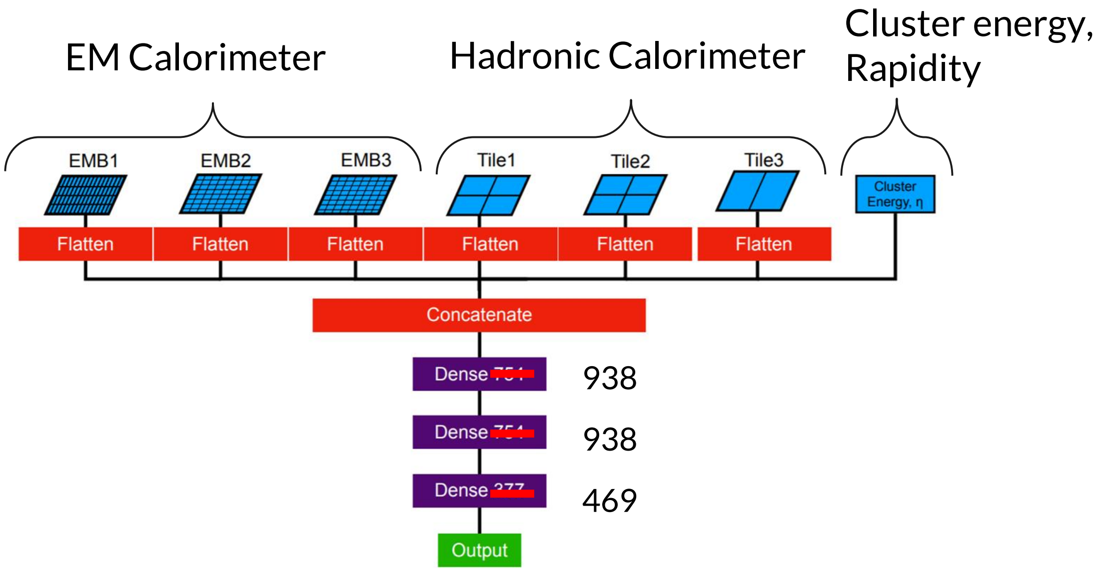
➔ Hadrons can deposit energy in matter through a series of successive interactions

  - Hadrons are relatively massive and lose their energy mainly through multiple nuclear collisions

➔ Atlas Hadronic Calorimeter is a sampling calorimeter using iron as absorber material and scintillating tiles as active material



DNN Regressor



Output : Predicted energy of the incident Pion

Models (DNN vs MDN)

- ➔ Dense layers have 938 neurons (number of cells + cluster energy)

DNN

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 938)	880782
dense_2 (Dense)	(None, 938)	880782
dense_3 (Dense)	(None, 469)	440391
dense_4 (Dense)	(None, 1)	470

Total params: 2,202,425
Trainable params: 2,202,425
Non-trainable params: 0

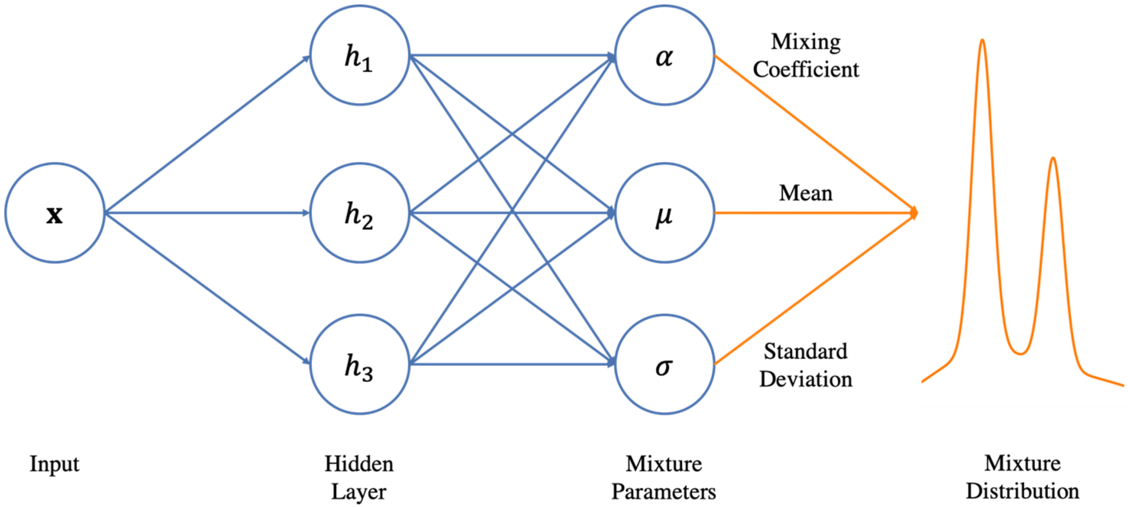
- ➔ MDN model adds additional MDN layers to the DNN

MDN

Layer (type)	Output Shape	Param #
dense_8 (Dense)	(None, 938)	880782
dense_9 (Dense)	(None, 938)	880782
dense_10 (Dense)	(None, 469)	440391
dense_11 (Dense)	(None, 3)	1410
mixture_normal_2 (MixtureNor	multiple	0

Total params: 2,203,365
Trainable params: 2,203,365
Non-trainable params: 0

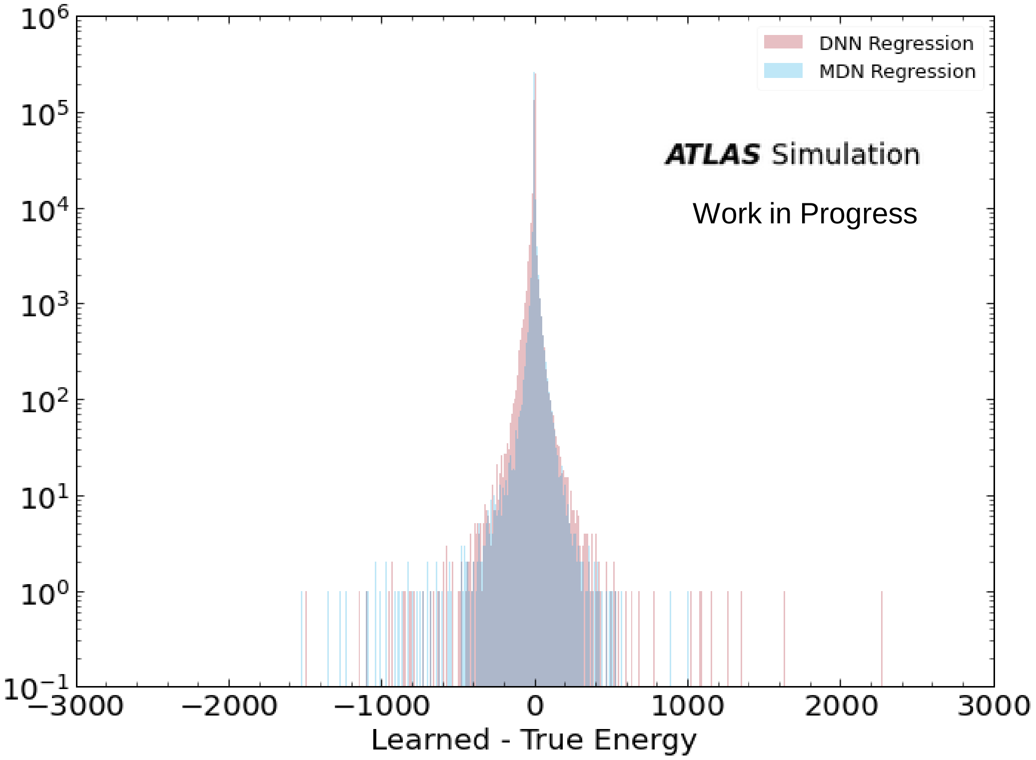
MDN Layer



Output : Predicted energy (mean) + predicted resolution (uncertainty) of the energy of the incident Pion

Results

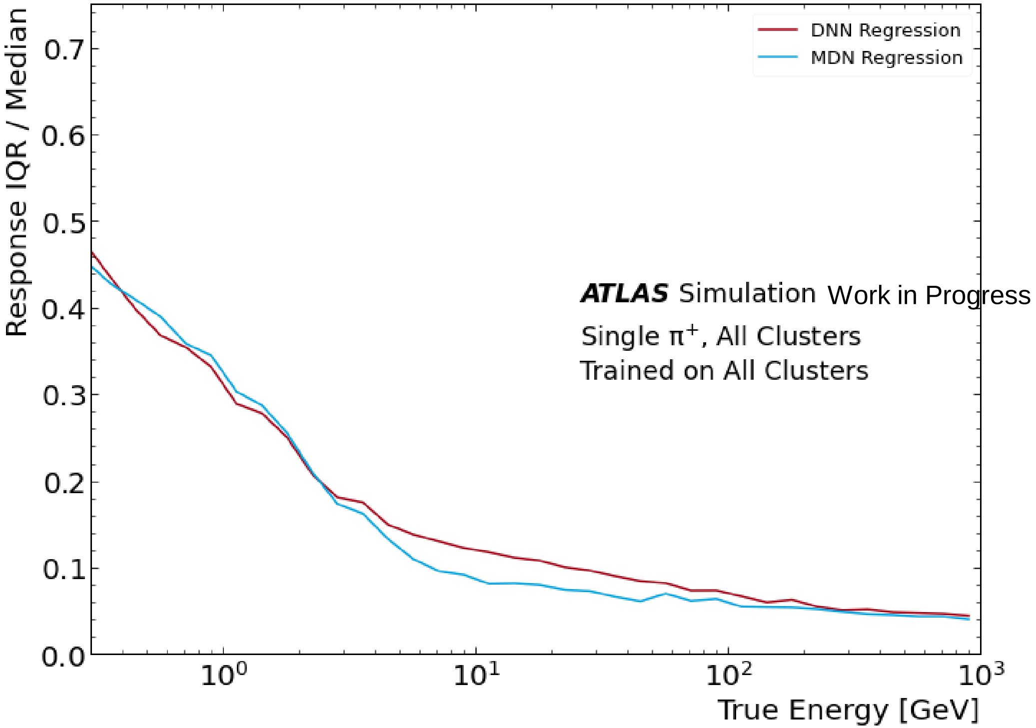
- ➔ Both DNN and MDN models have predicted - true energy in a normal curve shape centered around 0
- ➔ DNN seems to overpredict energy slightly more often



Results

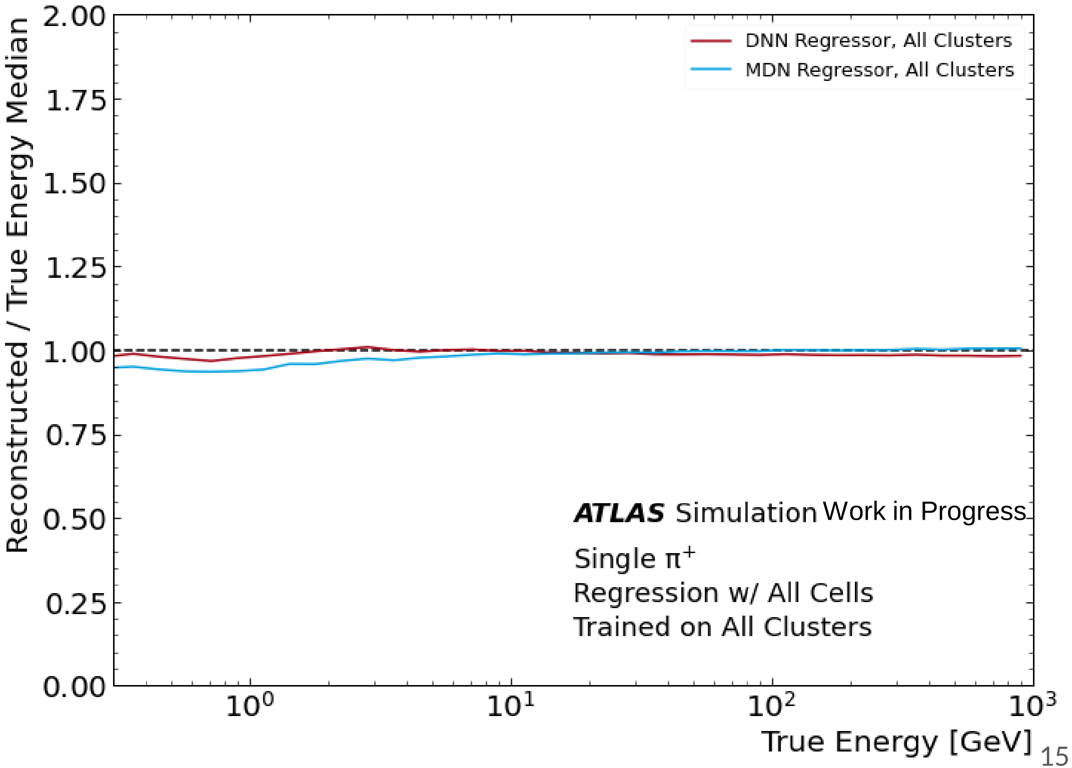
➔ Both DNN and MDN have a similar shape, where both have a much smaller Response IQR at higher energies

➔ The DNN and MDN perform similarly at lower energies



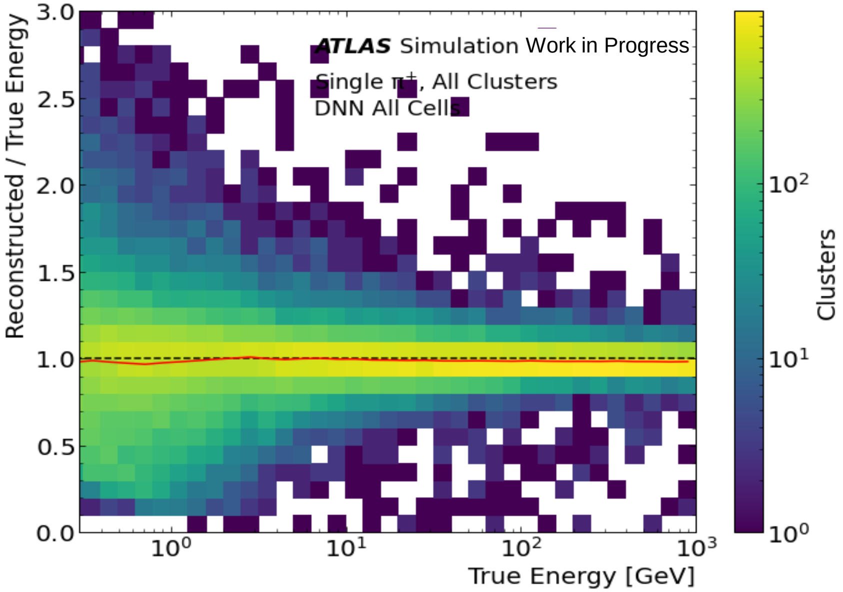
Results

- ➔ At lower energies <10 GeV, MDN performs much worse than at higher energies and underpredicts the true energy
- ➔ DNN also underpredicts at lower energies < 10 GeV, but performs better than the MDN

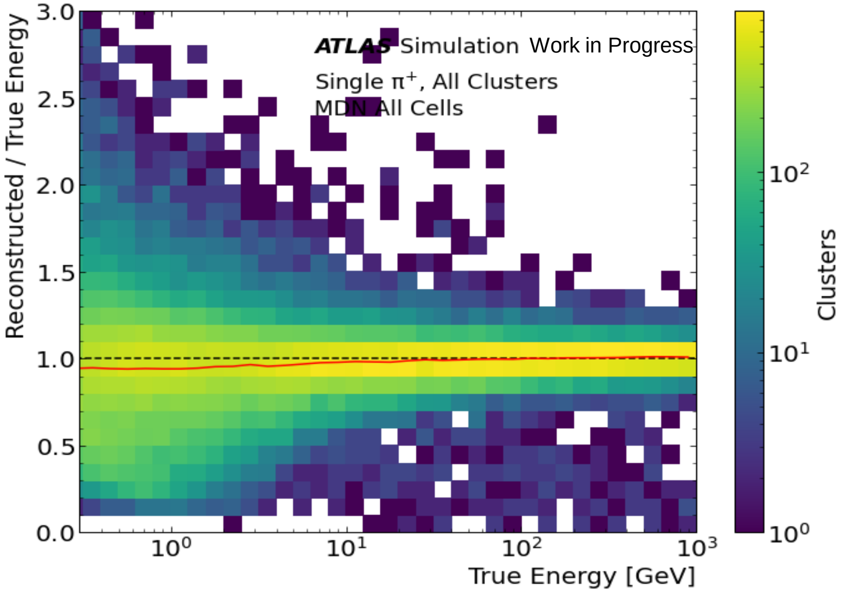


Results

DNN



MDN



Conclusion + Future Possibilities

The MDN very slightly underperforms compared to DNN, but predicts previously unquantified information (resolution of the pion energy per cluster)

Future Improvements:

1. Hyperparameter search to get the most optimal results and compare again
2. Compare predicted resolutions to previous measurements to understand model accuracy
3. Add more components to MDN layer to outperform DNN model