



UNIVERSIDAD
COMPLUTENSE
MADRID



CTAO | LST
COLLABORATION

Deep Neural Network compression for low-latency event characterisation in imaging atmospheric Cherenkov Telescopes

Dafne Martín Domínguez (dafmar02@ucm.es)

D. Nieto

UCM-GAE

10th December 2025



Grant PDC2023-145839-I00 funded by:



Contextualisation: Atmospheric showers

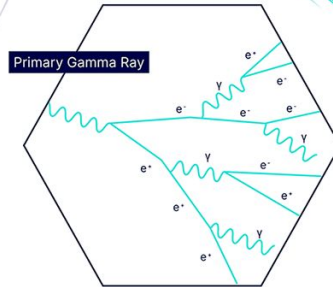


Very energetic
astrophysical
phenomena

Gamma Ray

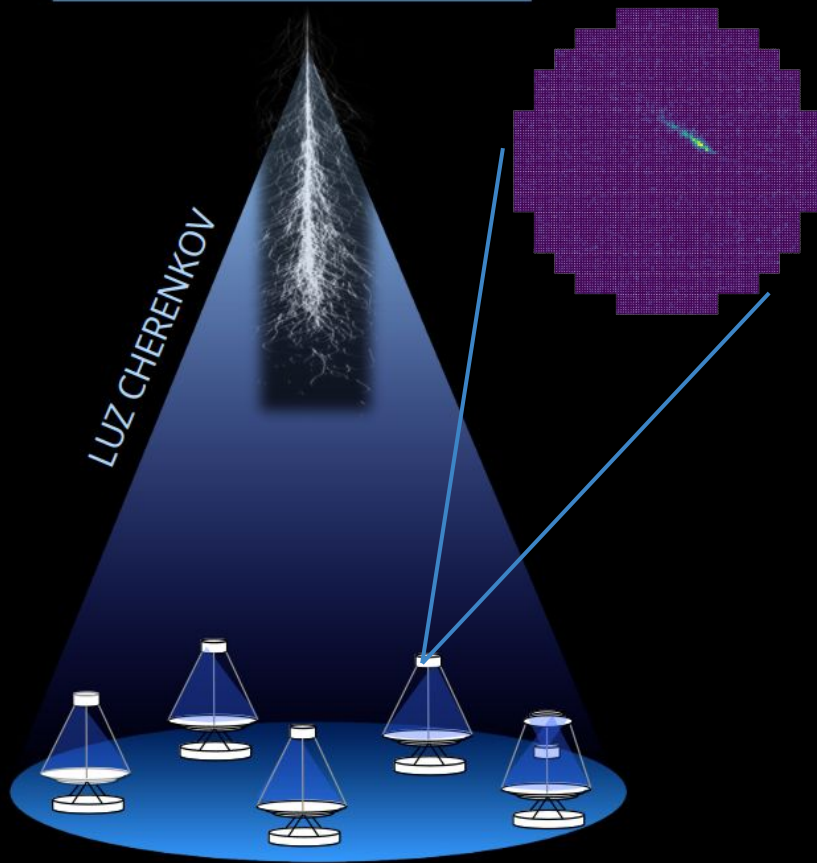
Atmosphere

Air Shower



gamma ray interacts with the atmosphere

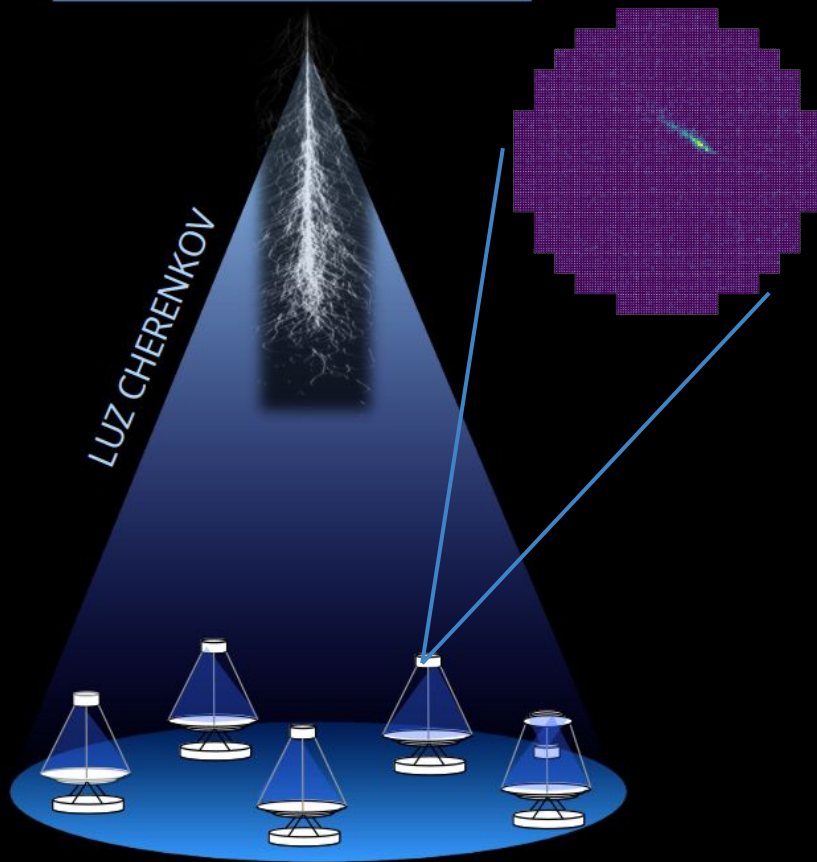
ATMOSPHERE



IACT technique

- **Trigger** (see Maria Molina's talk)
 - Background fluctuations
 - Atmospheric showers
- **Reconstruction**
 - Particle type
 - Direction
 - Energy

ATMOSPHERE



IACT technique

- **Trigger** (see Maria Molina's talk)
 - Background fluctuations
 - Atmospheric showers
- **Reconstruction**
 - Particle type
 - Direction
 - Energy
- **Challenge: *Low latency***

Cherenkov Telescope Array Observatory

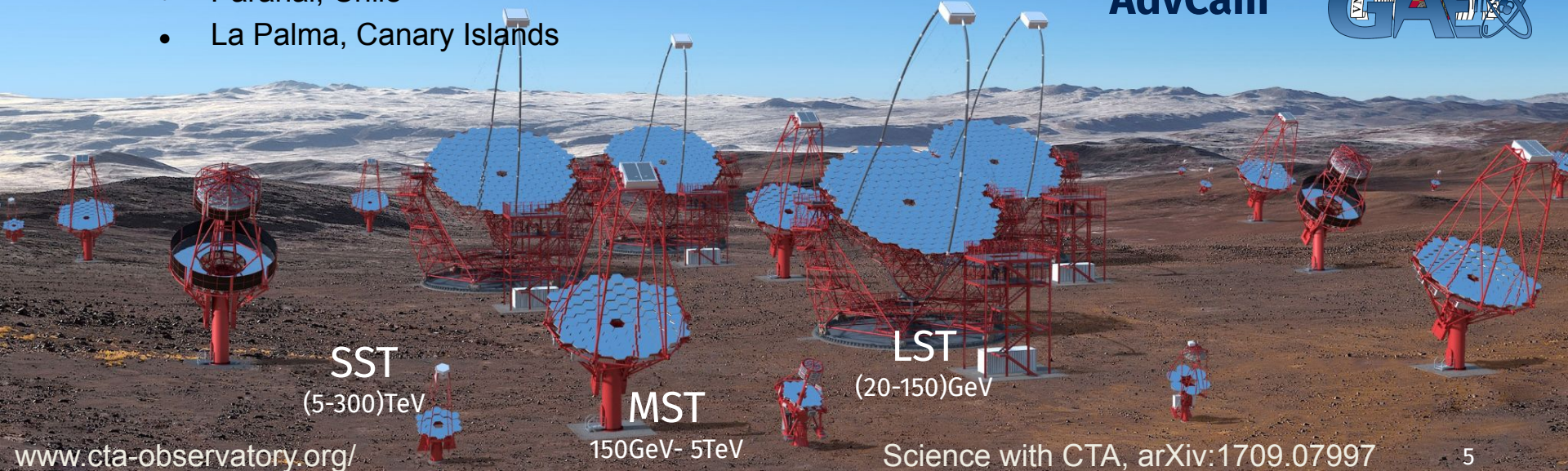
Cherenkov Telescope Matrix (CTAO)

- Earth-base observatory for gamma astronomy.
- Two sites:
 - Paranal, Chile
 - La Palma, Canary Islands

CTAO

LST
COLLABORATION

AdvCam



SST

(5-300)TeV

MST

150GeV- 5TeV

LST

(20-150)GeV

Hypothesis and objectives: Model compression

Explore model compression for ResNet-36 (DL)

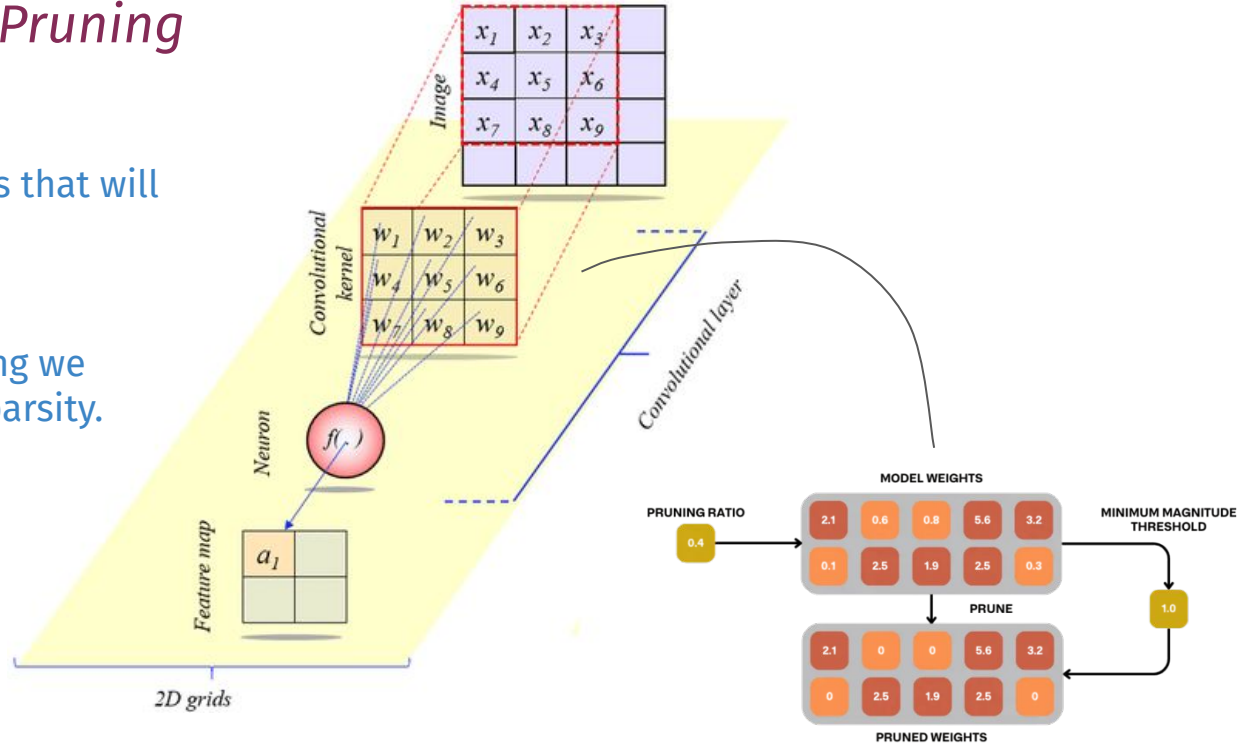
Objectives

- Identify and explore model compression techniques
- Implement:
 - Pruning
- Evaluate degradation in performance
 - Metrics
 - IRFs
- Evaluate model size reduction

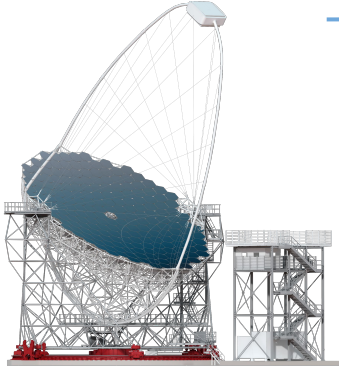
Methodology

Magnitude-based Pruning

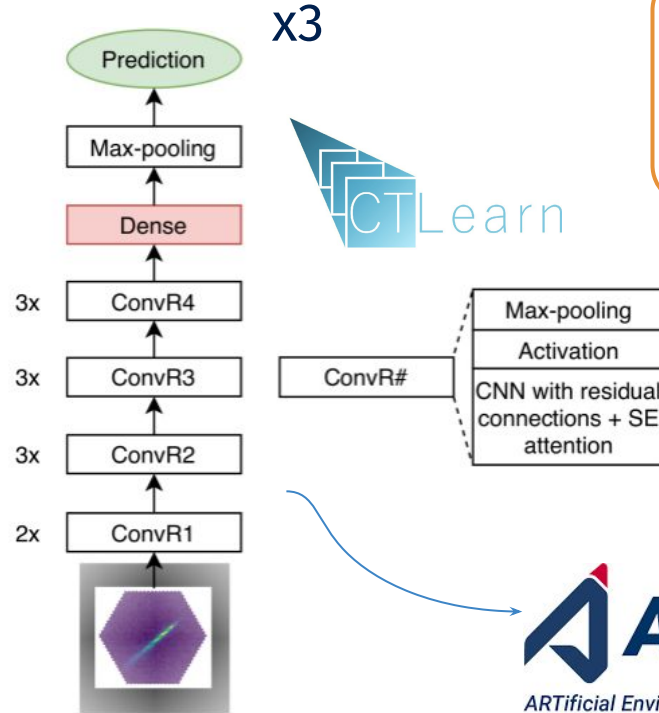
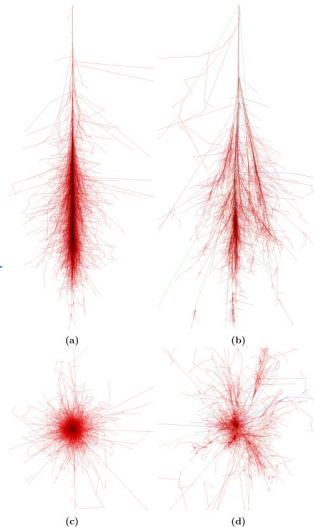
- Sparsity = How much?
Percentage of weights that will be zeroed out.
- Schedule = How?
How quickly in training we obtain the desired sparsity.
 - Constant
 - Polynomial



Methodology



Implement magnitude based pruning



Evaluation:

- Metrics
- IRFs

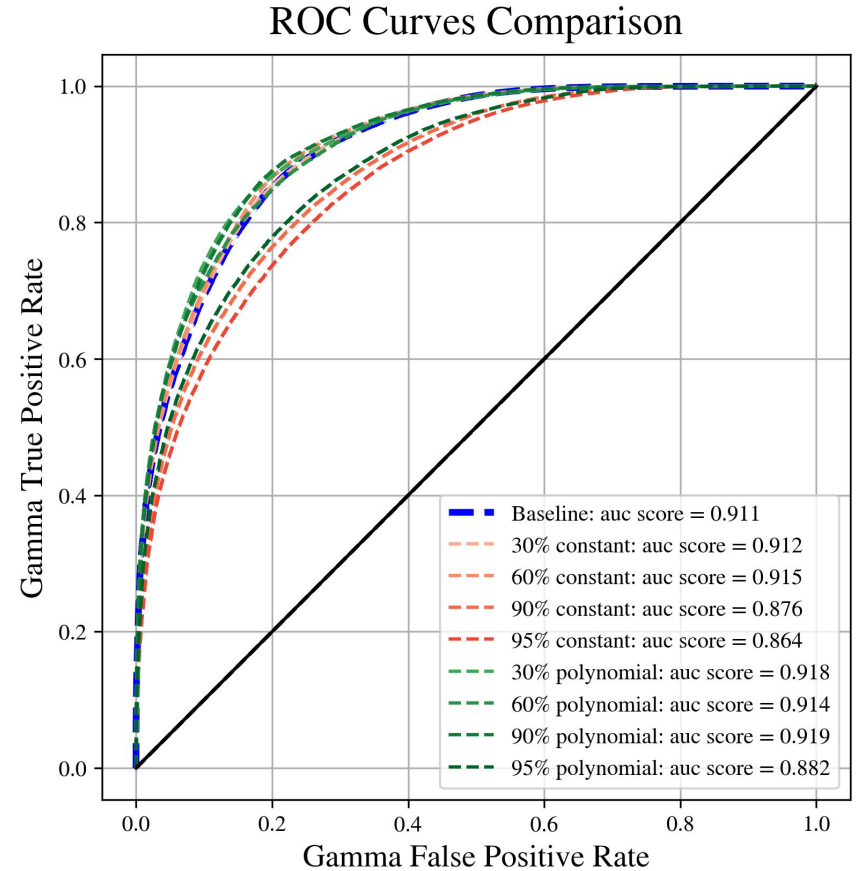
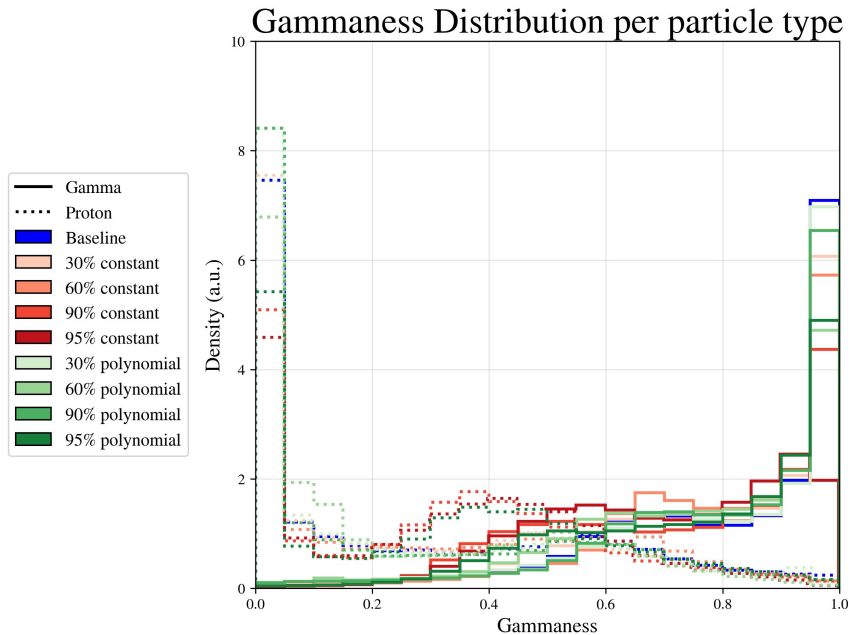


Results: Classification task

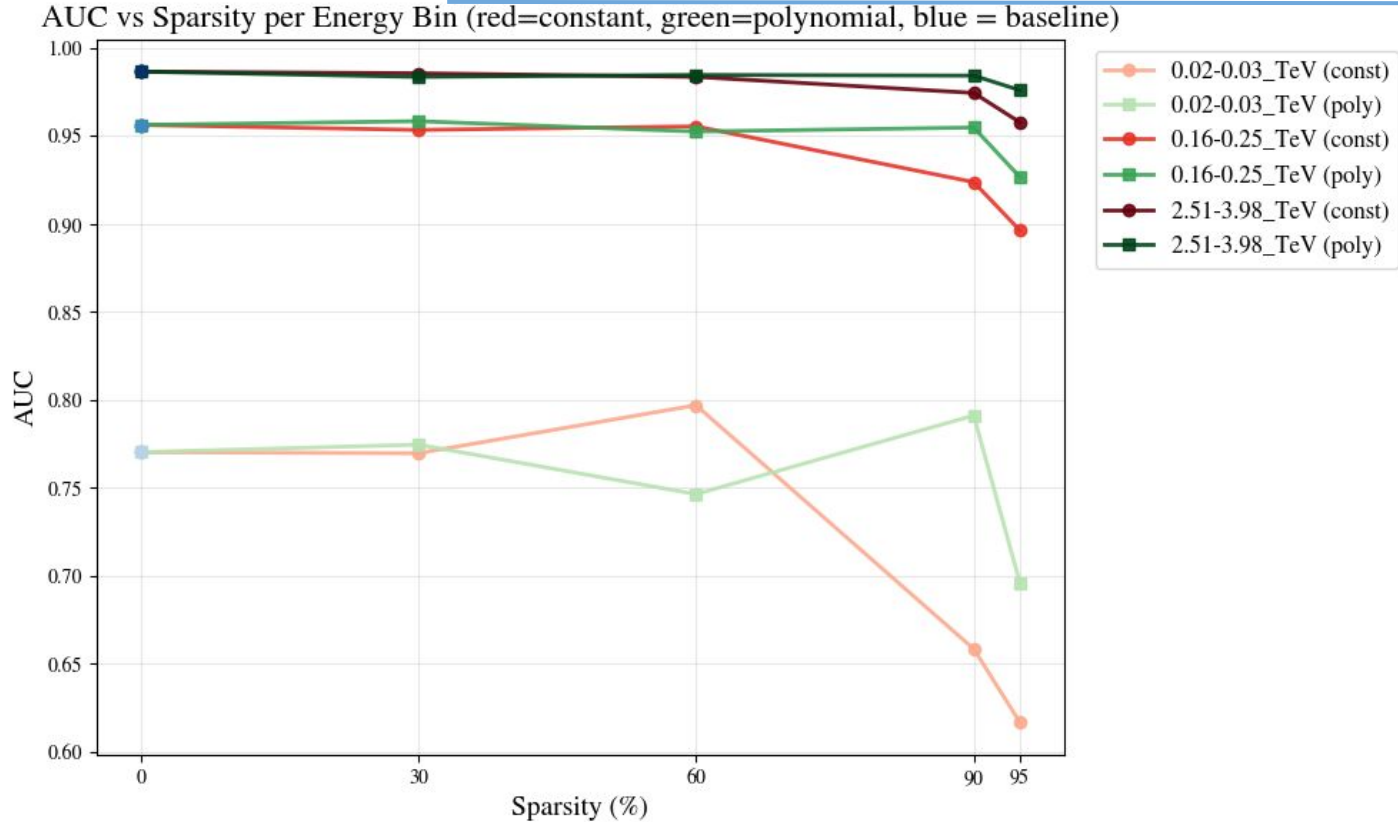
- Compare with Baseline
- *Gammaness* histograms
- ROC curves: performance of a binary classifier as a function of gammaness
- Area under curve (AUC)



Results: Classification task



Results: Classification task

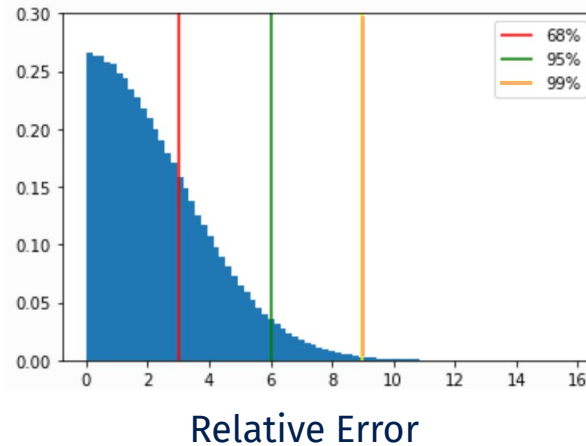


Results: Regression task

Resolution Curves

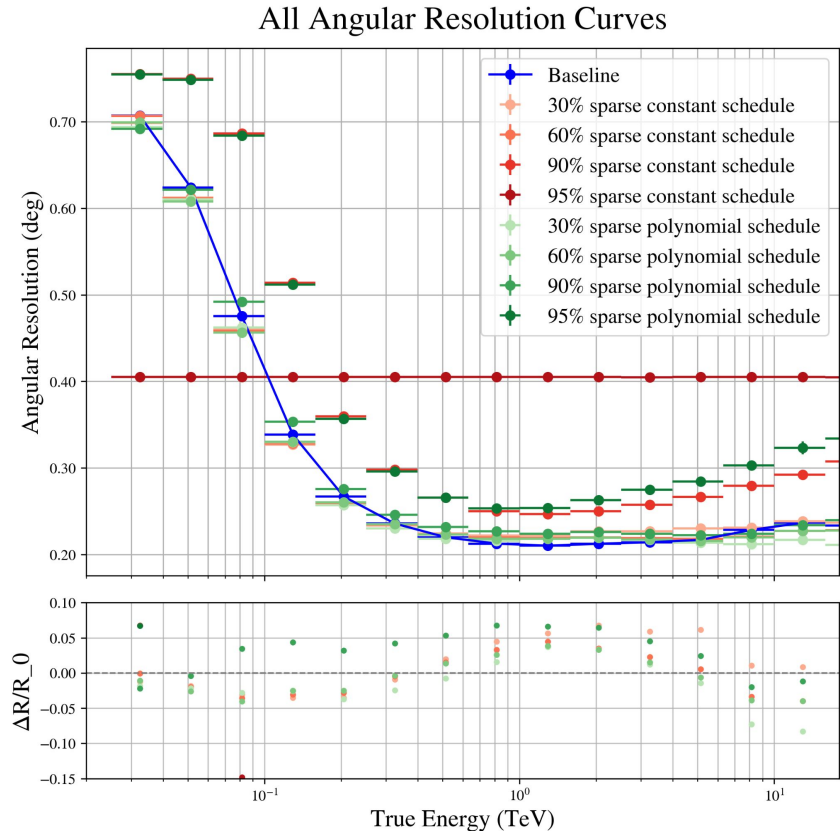
Relative error value under which 68% of the reconstructed events produced by a gamma ray are contained.

$$\text{Relative Error} = \frac{|\text{Predicted value} - \text{Real value}|}{\text{Real value}}$$

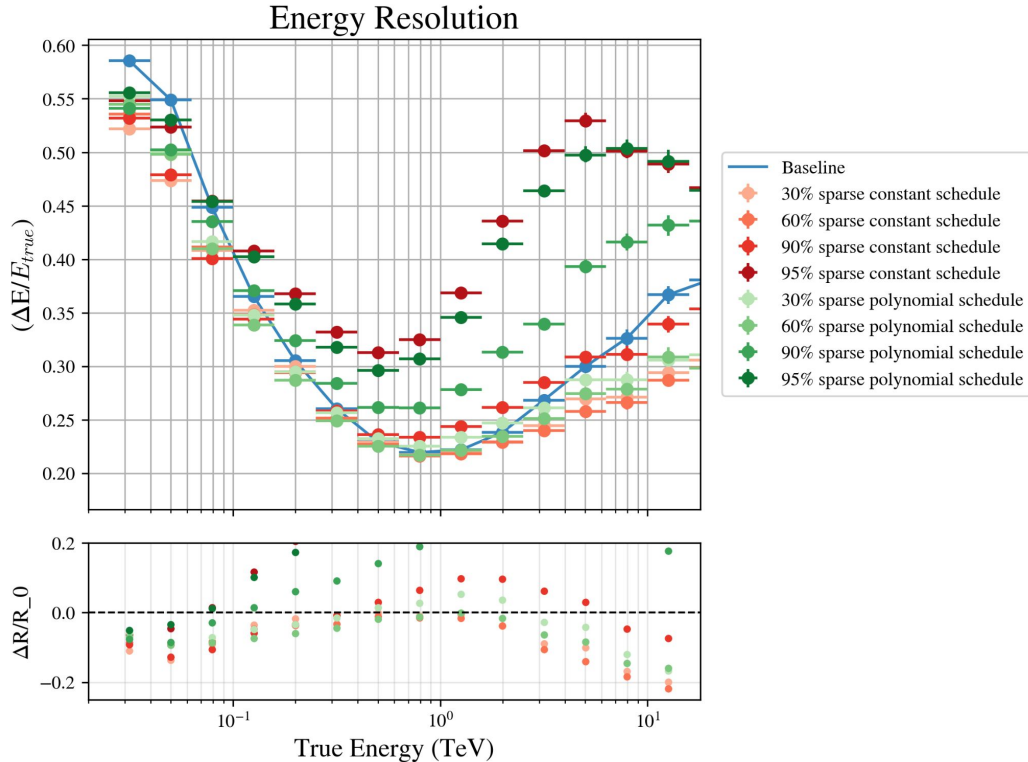


Results: Regression task

- Polynomial: Similar to baseline until sparsities up to 90%
- Polynomial outperforms constant schedule
- Possible improvement in some energy ranges
- Relative deviation with respect to baseline

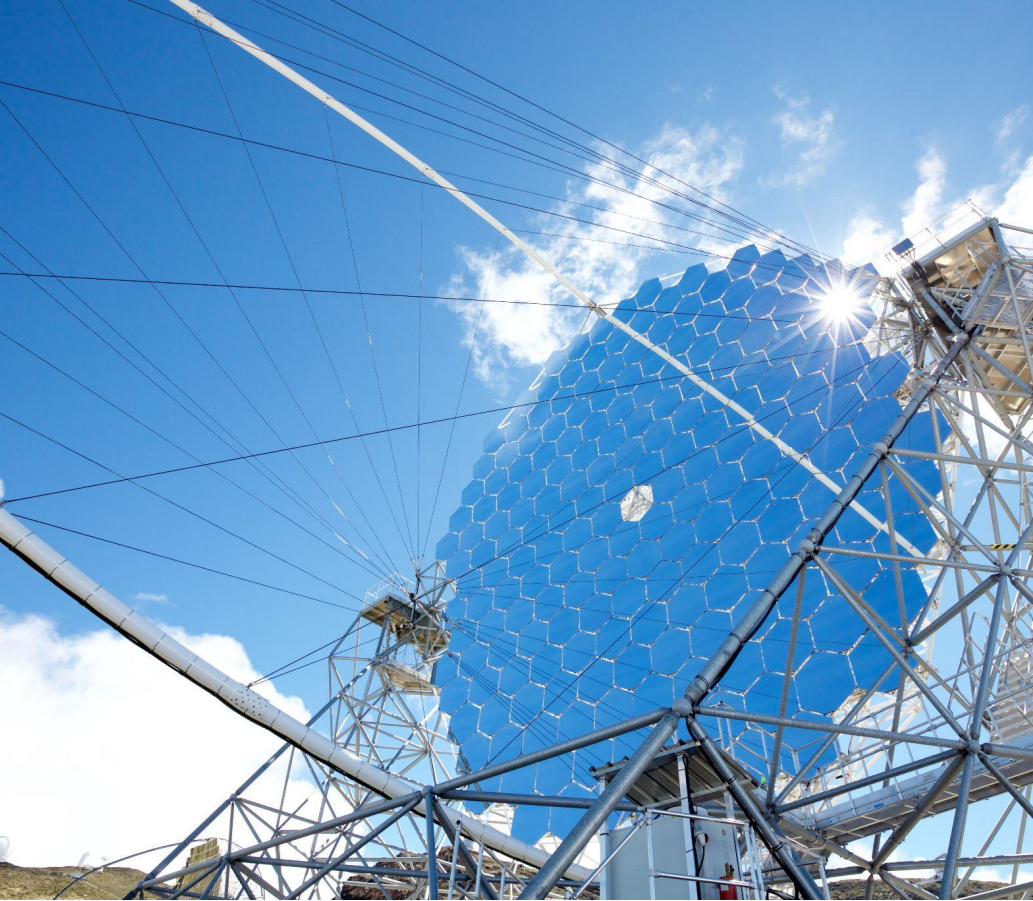


Results: Regression task



- Similar to baseline until sparsities up to 90% de sparsity
- Possible improvement in some energy ranges
- Relative deviation with respect to baseline

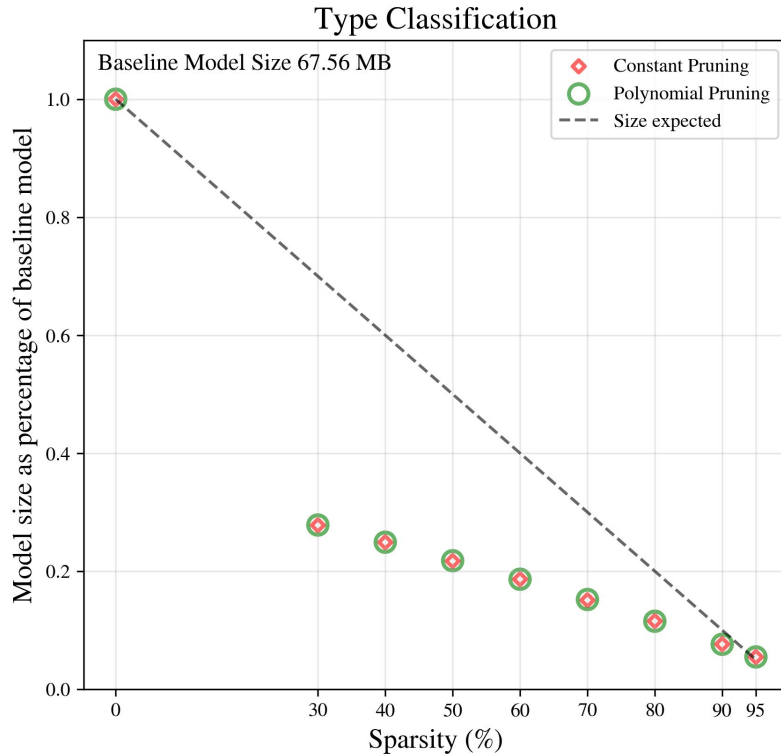
Results: Model size



Evaluate the size of the model as a function of sparsity

Results: Reduction in size

Model Size vs Sparsity for Different Tasks



- Model sizes after compression to .zip
- Smaller size than linear reduction
- 30% sparsity: 28% of original size
- 95% sparsity: 5% of original size

Conclusions & Future Work

Conclusions

- **Polynomial sparsity** outperforms **constant sparsity**
- **Classification: polynomial schedule** without significant degradation up to 90-95%.
- **Energy polynomial schedule** without significant degradation at 90%.
- **Direction: polynomial schedule** without significant degradation up to 90%.
- Significant reduction in model **size**

Ongoing Work

- NSB vs. cosmics classification
- L1 on flowers
- Use simpler CNN model
- Sparsity-aware quantisation implementation

Acknowledgements



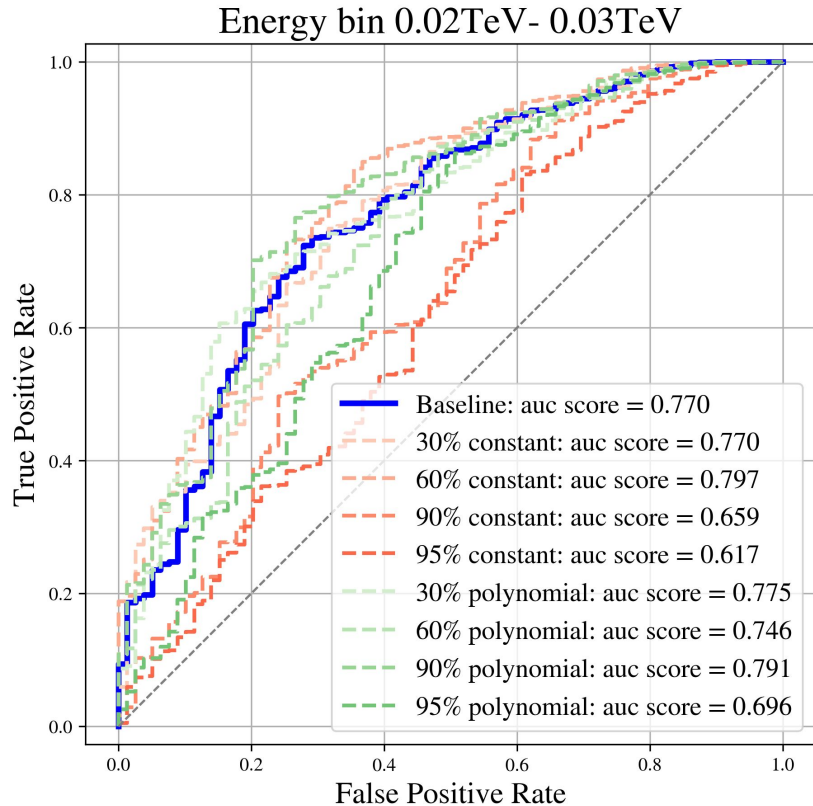
The research here presented has been partially supported by the MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR, under grant PDC2023-145839-I00



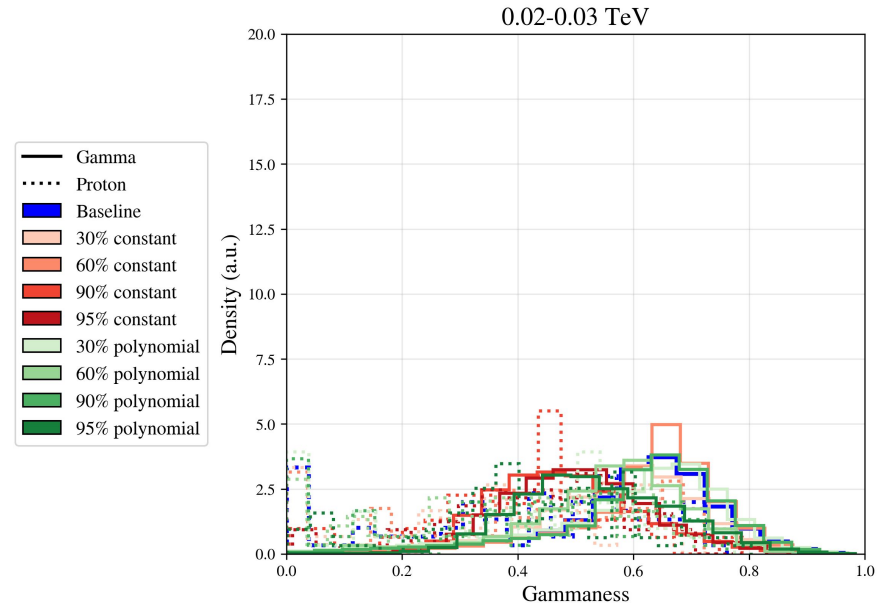
The author(s) gratefully acknowledges the computer resources at Artemisa, funded by the European Union ERDF and Comunitat Valenciana as well as the technical support provided by the Instituto de Física Corpuscular, IFIC (CSIC-UV)



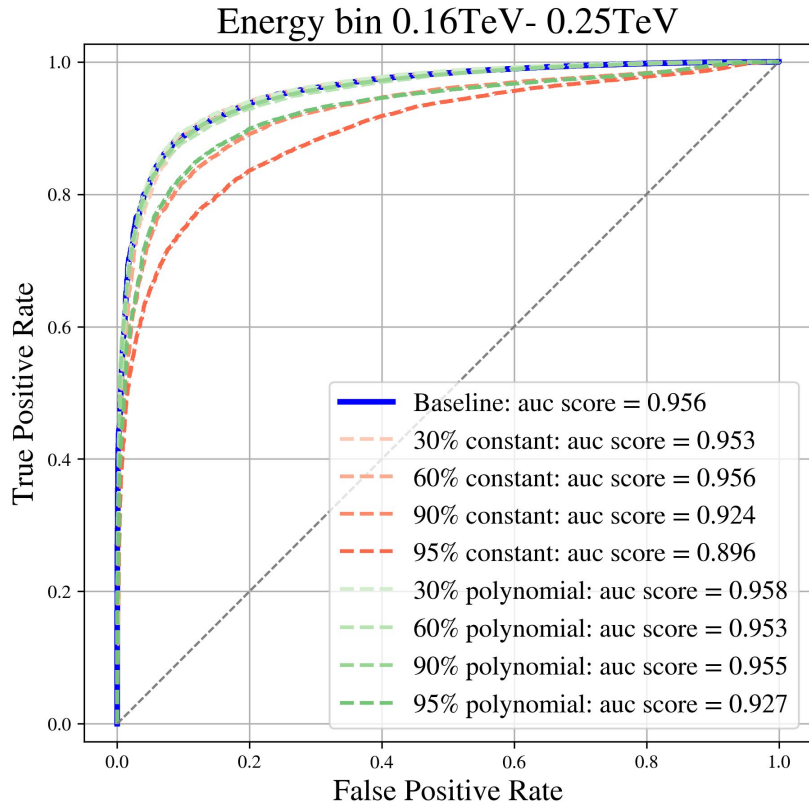
Results: Classification task



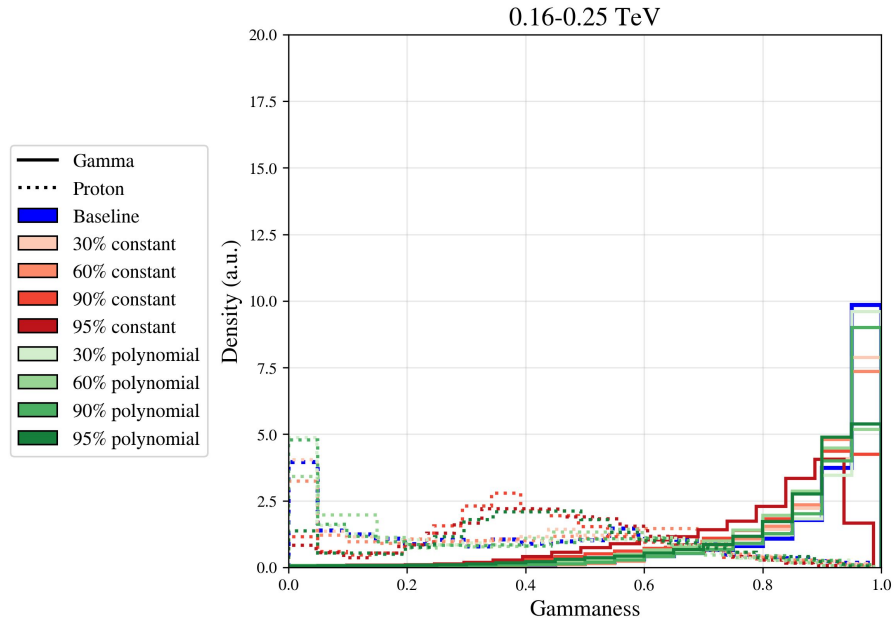
Gammaness Distribution Comparison per Energy Bin



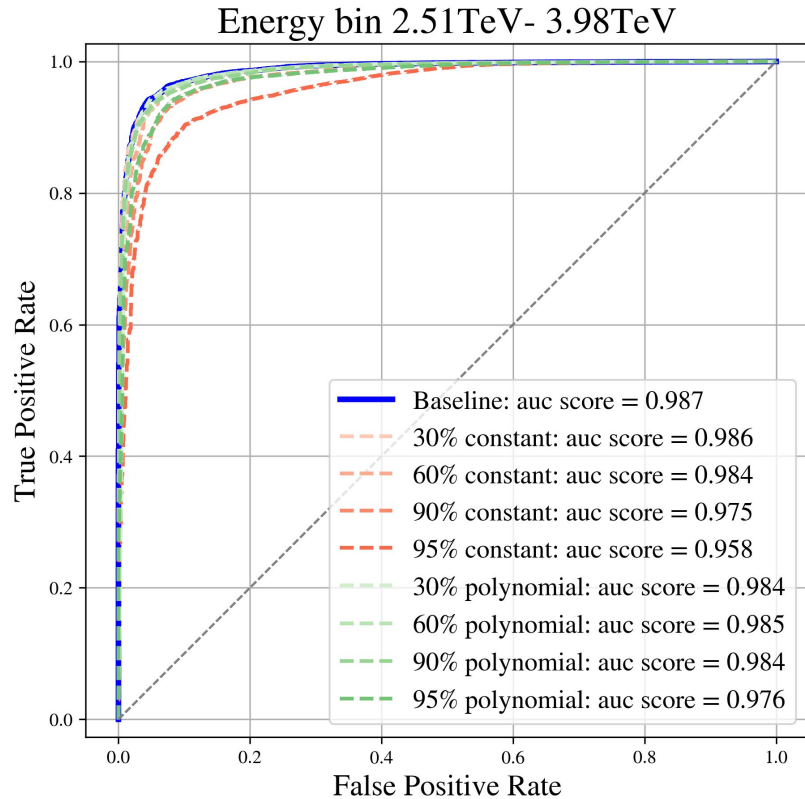
Results: Classification task



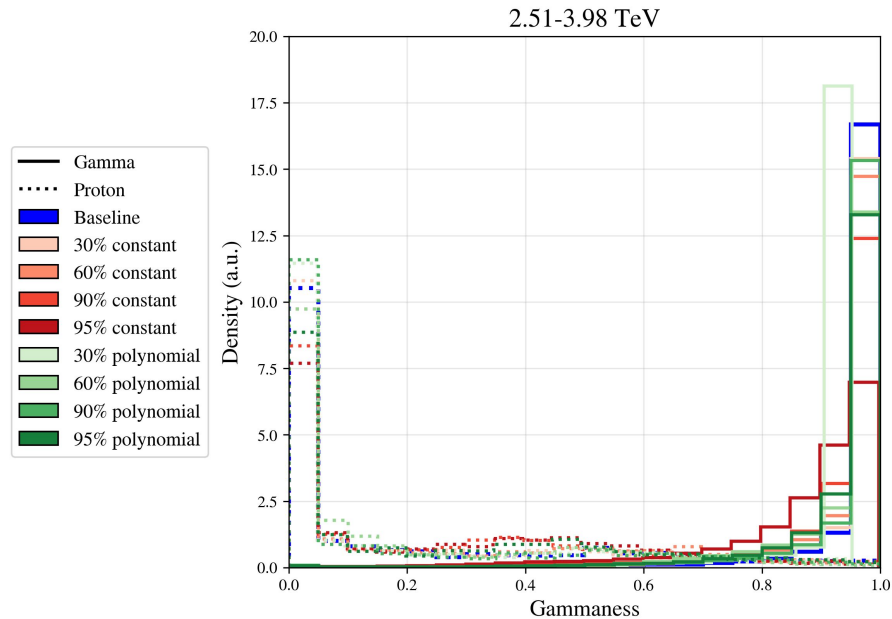
Gammaness Distribution Comparison per Energy Bin



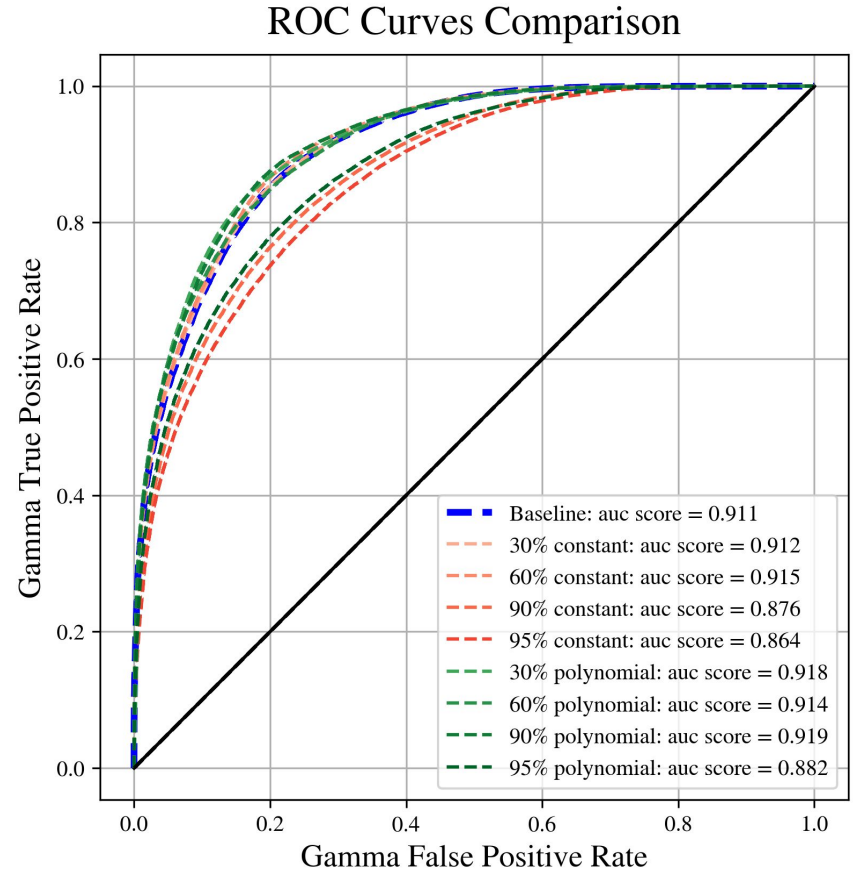
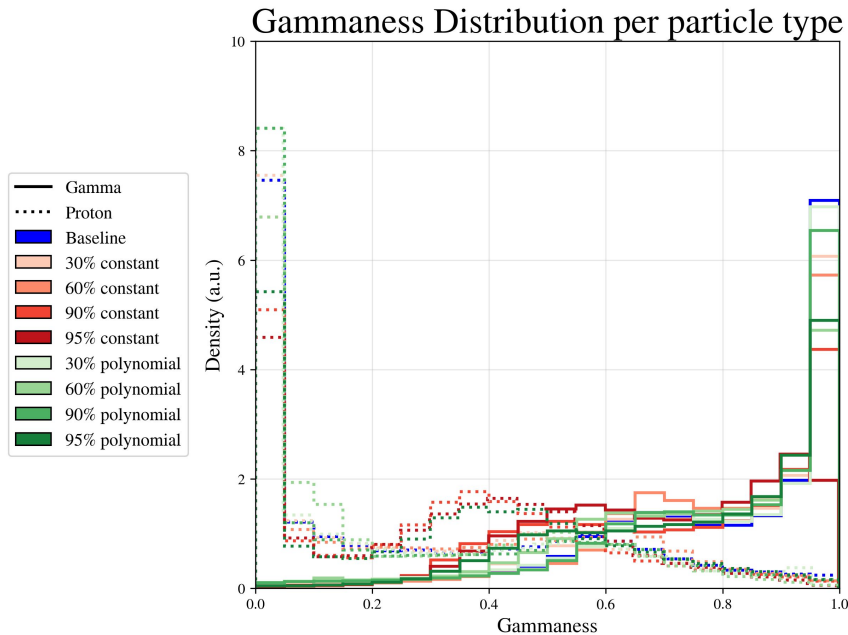
Results: Classification task



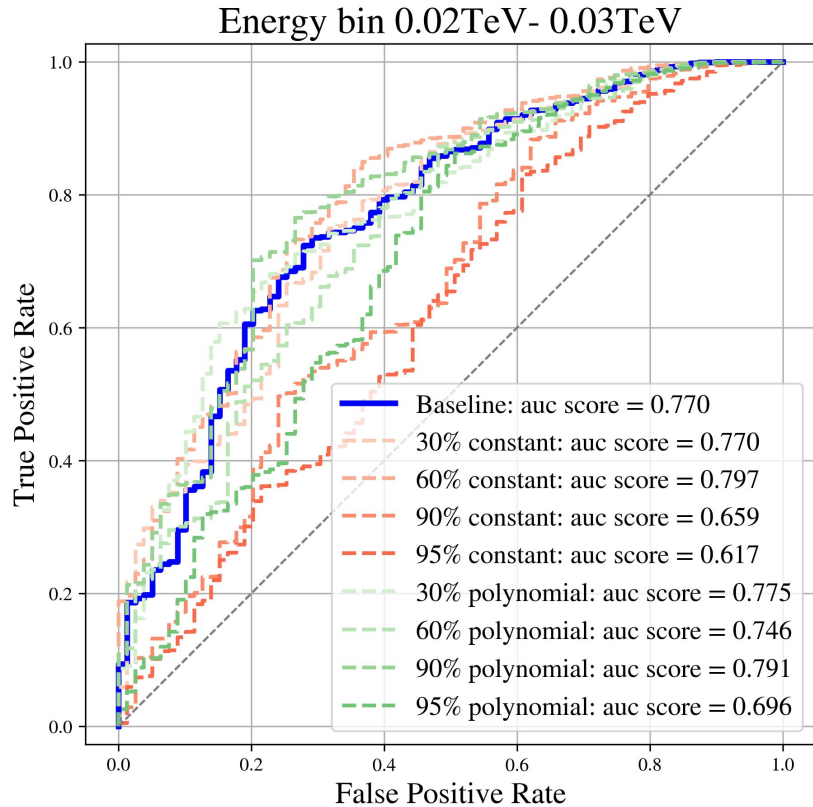
Gammaness Distribution Comparison per Energy Bin



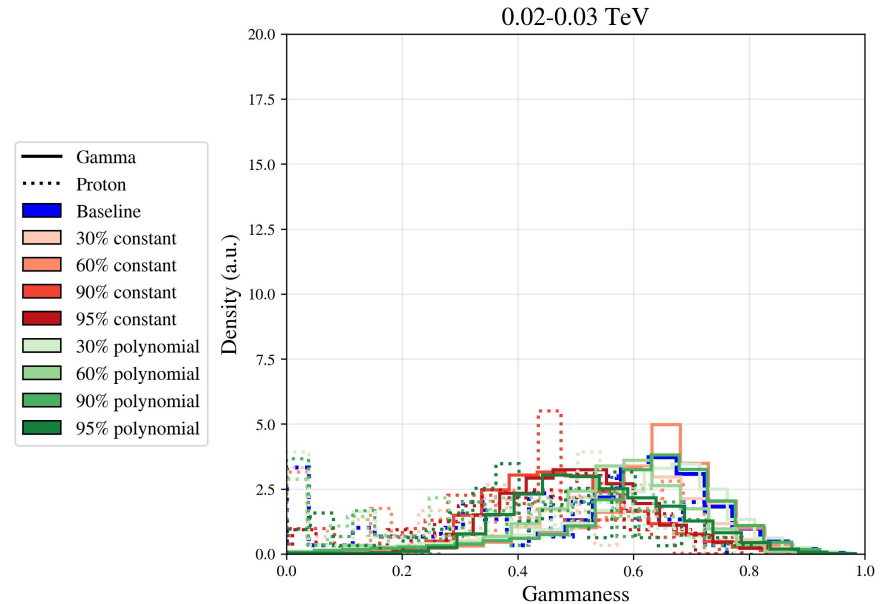
Results: Classification task



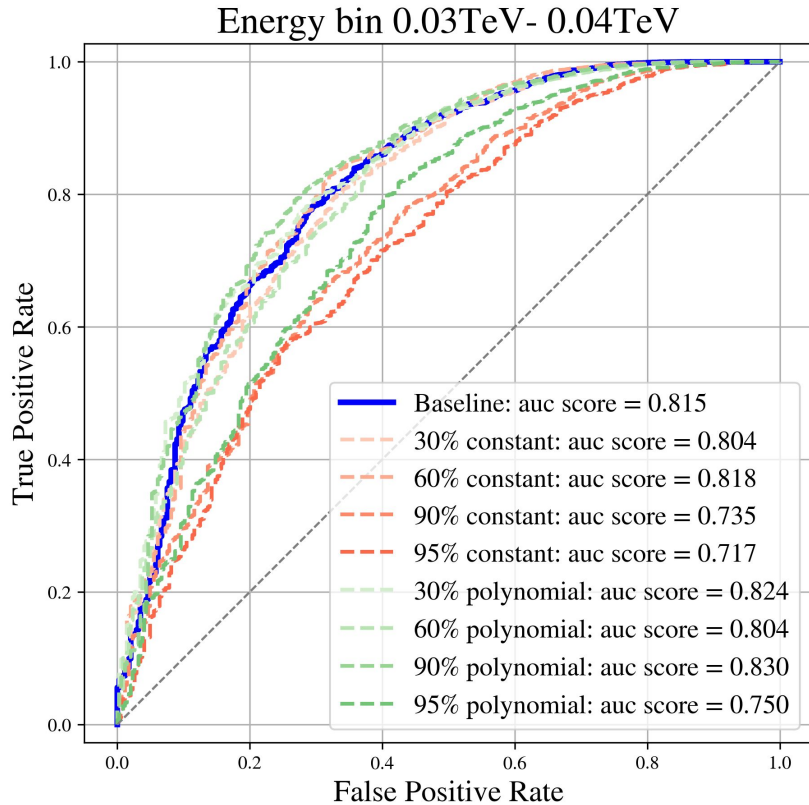
Results: Classification task



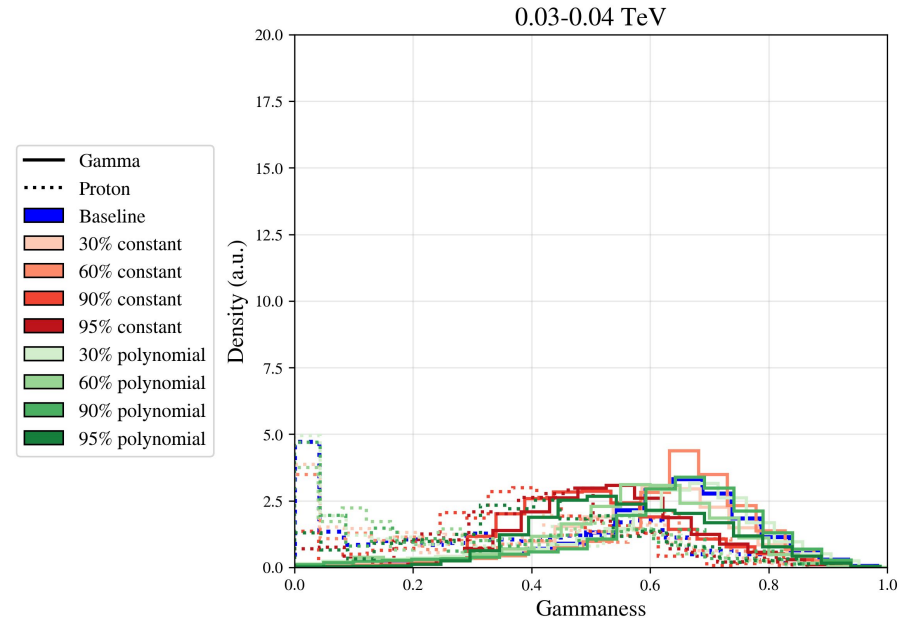
Gammaness Distribution Comparison per Energy Bin



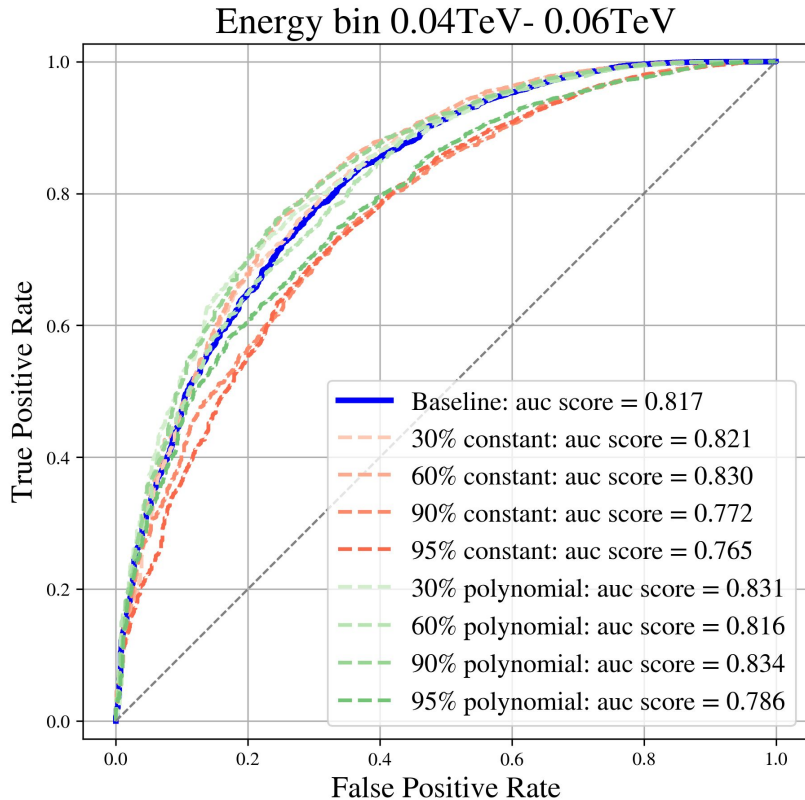
Results: Classification task



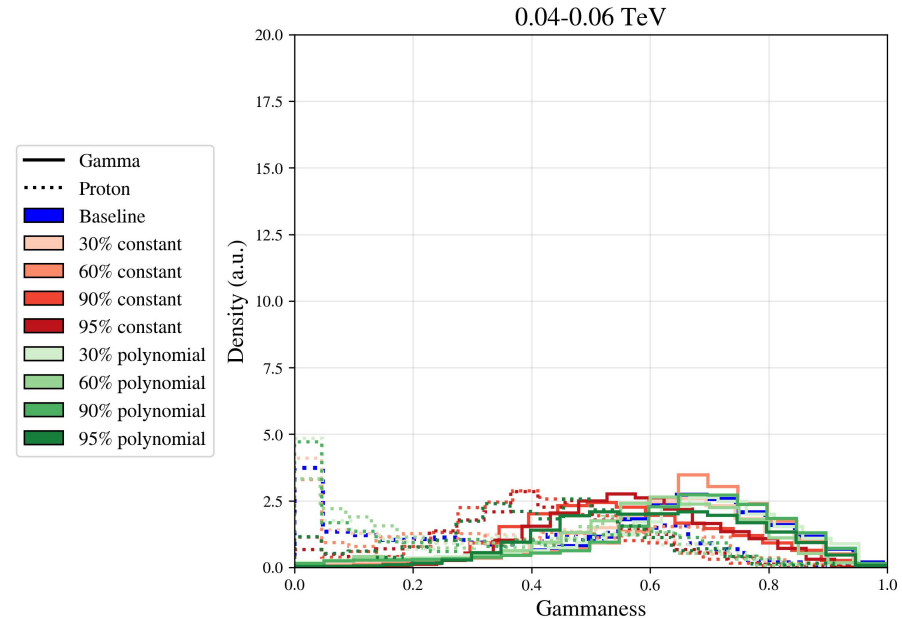
Gammaness Distribution Comparison per Energy Bin



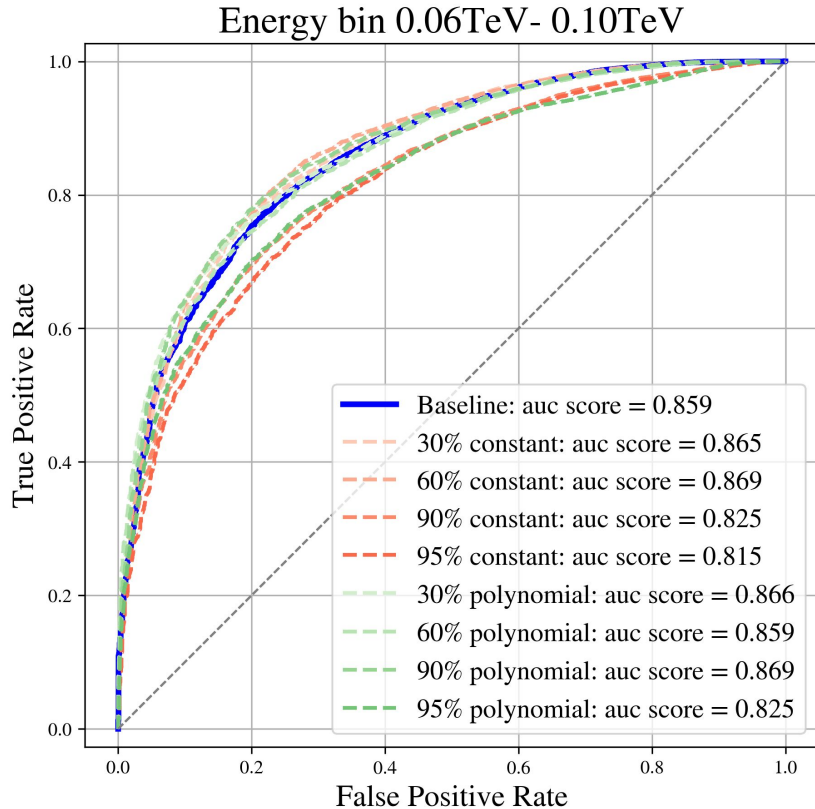
Results: Classification task



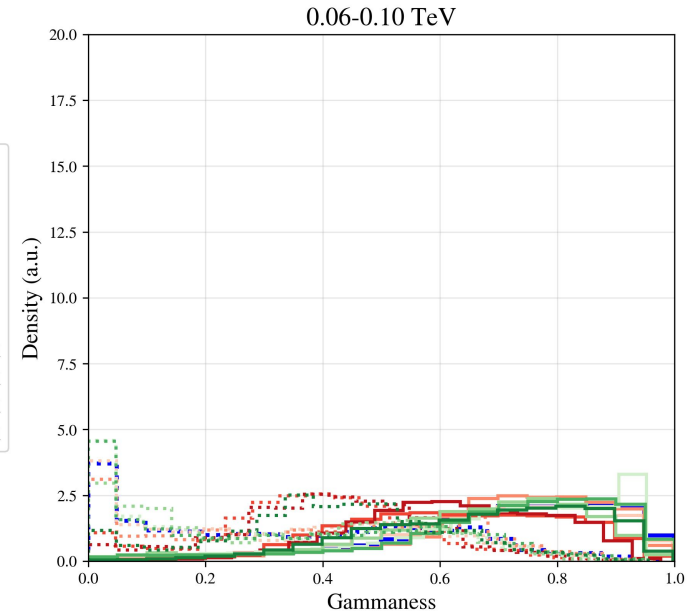
Gammaness Distribution Comparison per Energy Bin



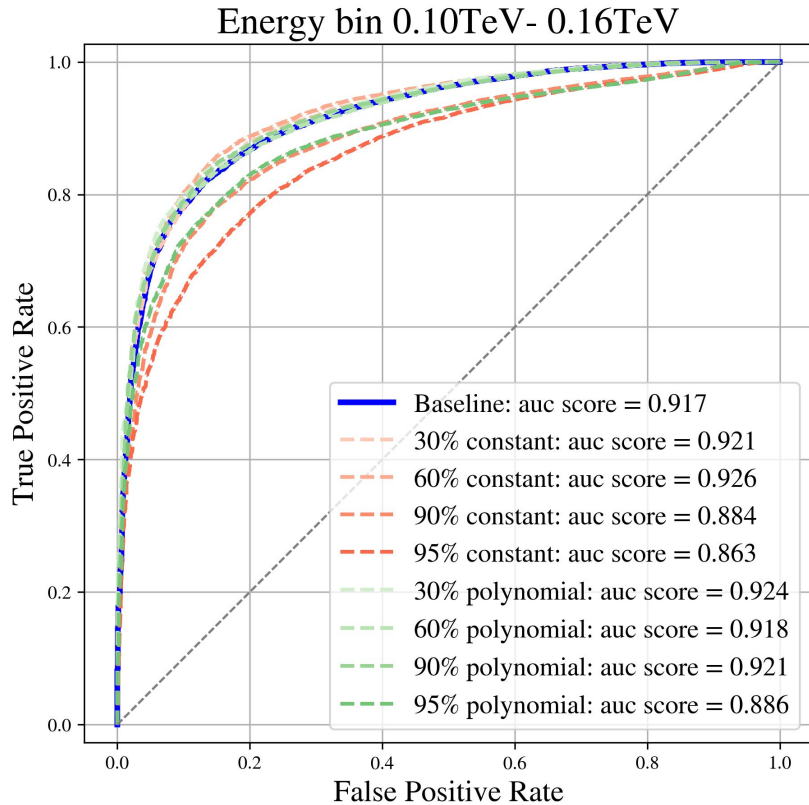
Results: Classification task



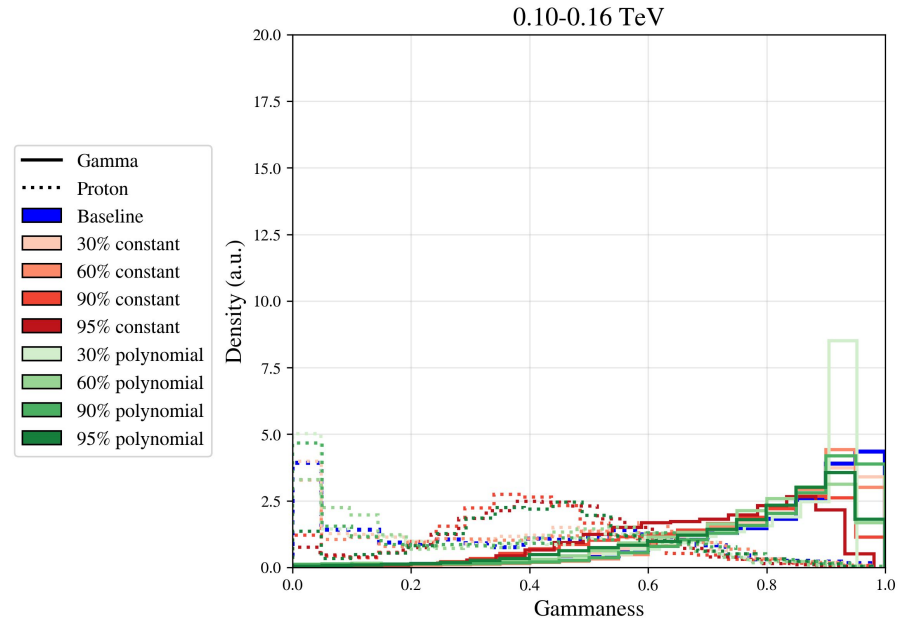
Gammaness Distribution Comparison per Energy Bin



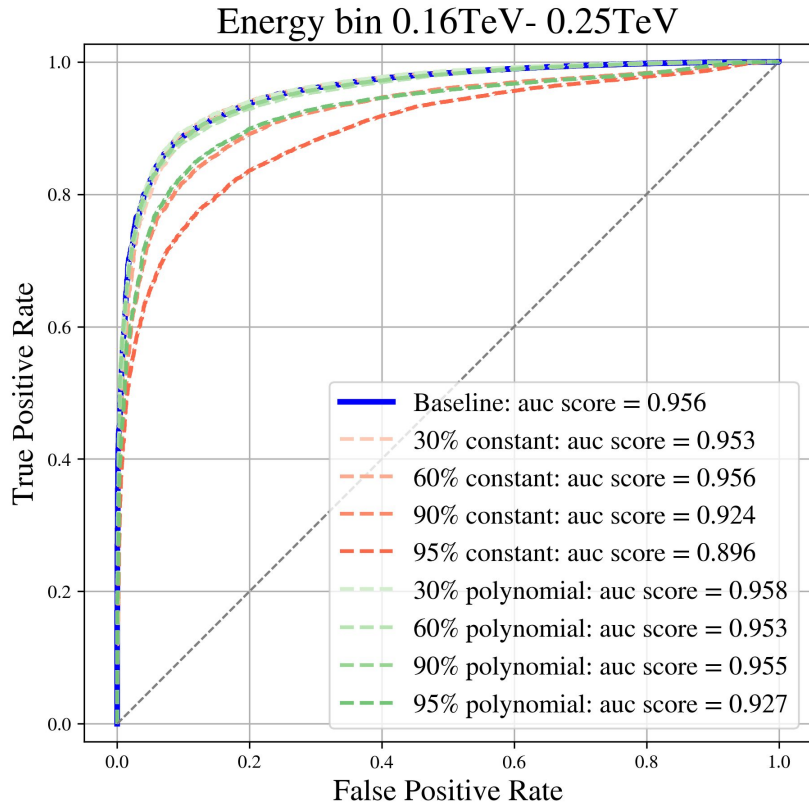
Results: Classification task



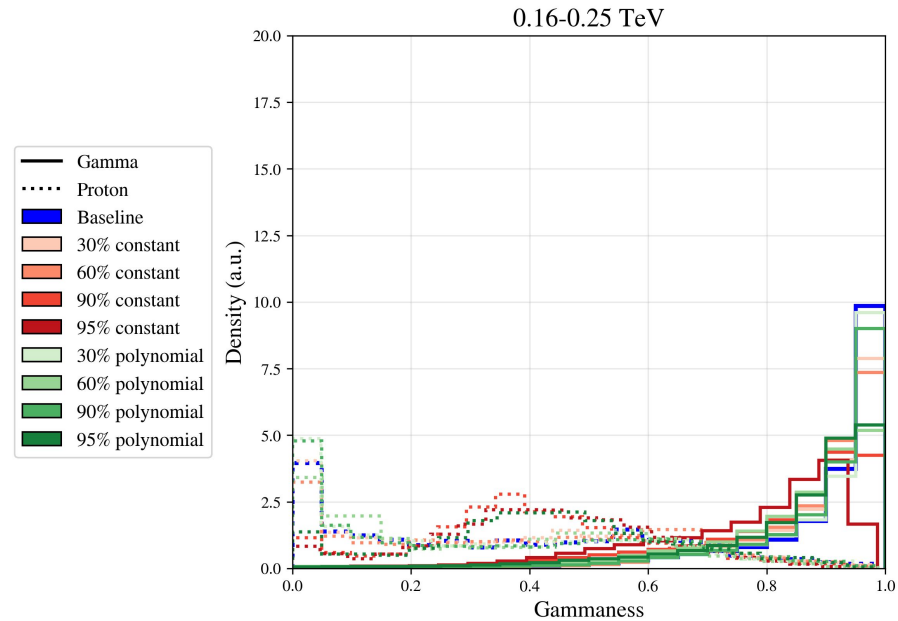
Gammaness Distribution Comparison per Energy Bin



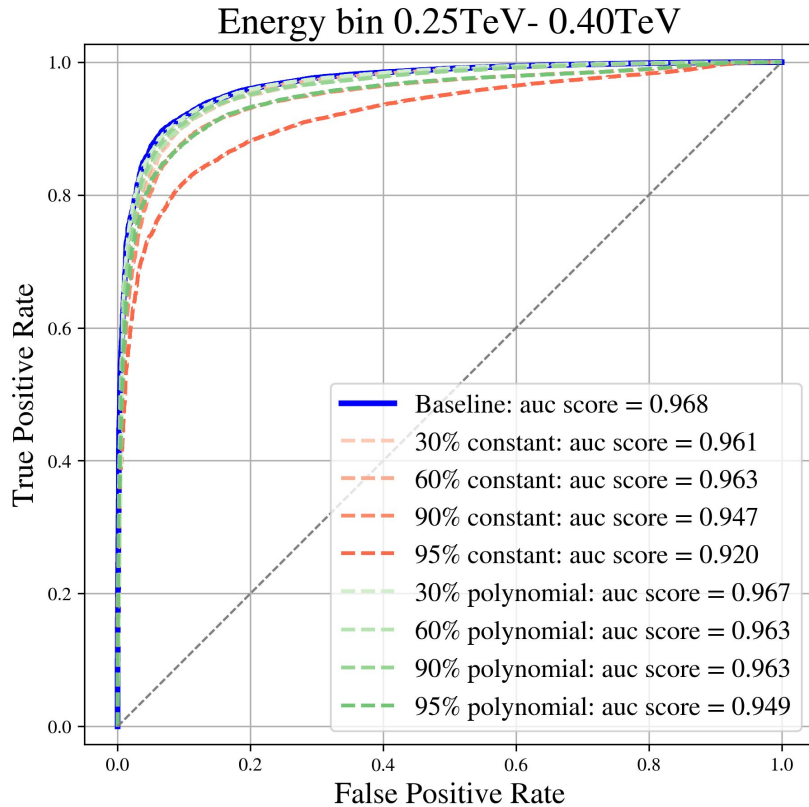
Results: Classification task



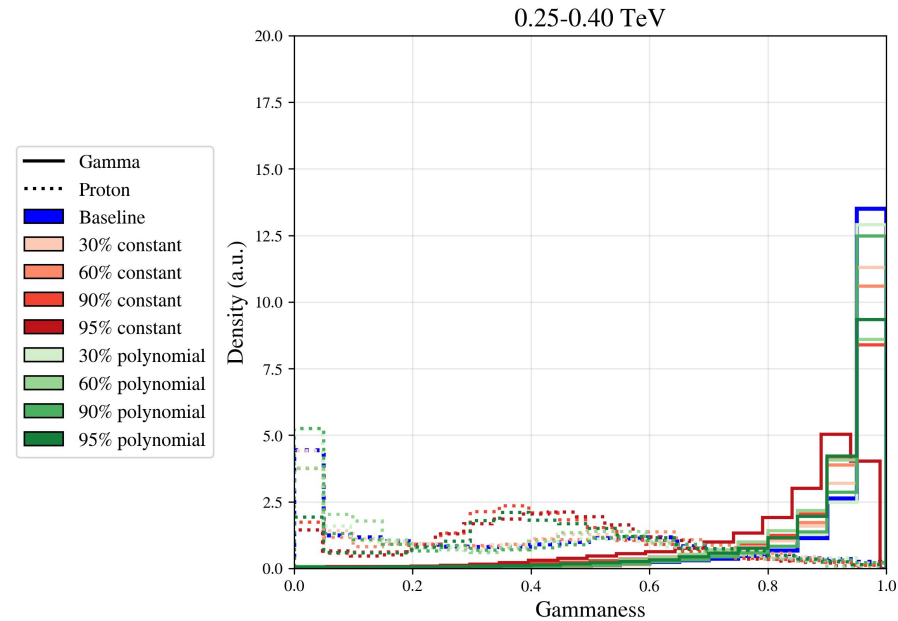
Gammaness Distribution Comparison per Energy Bin



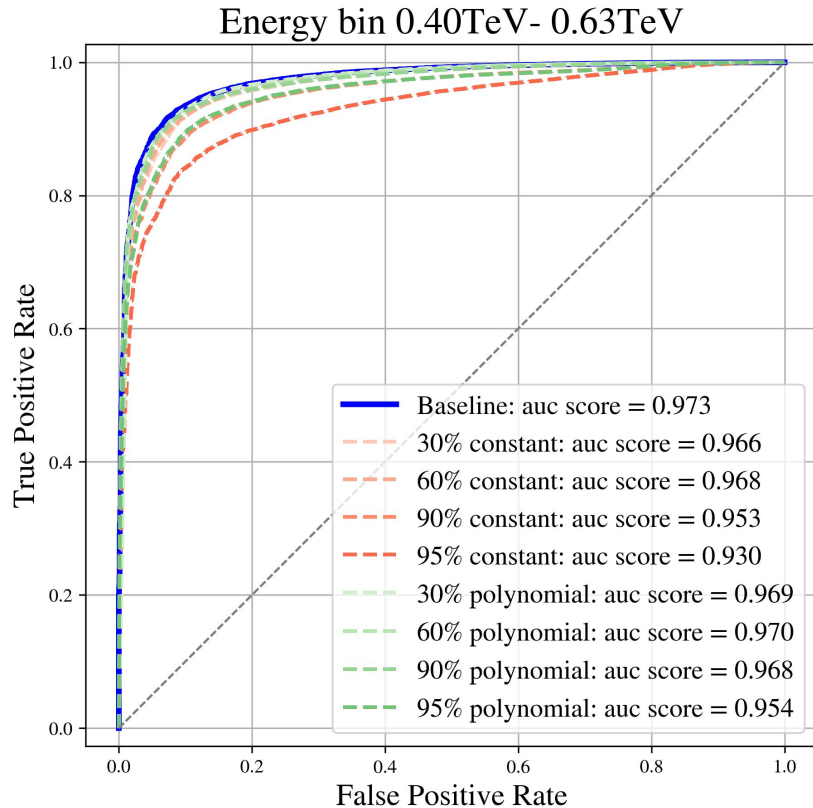
Results: Classification task



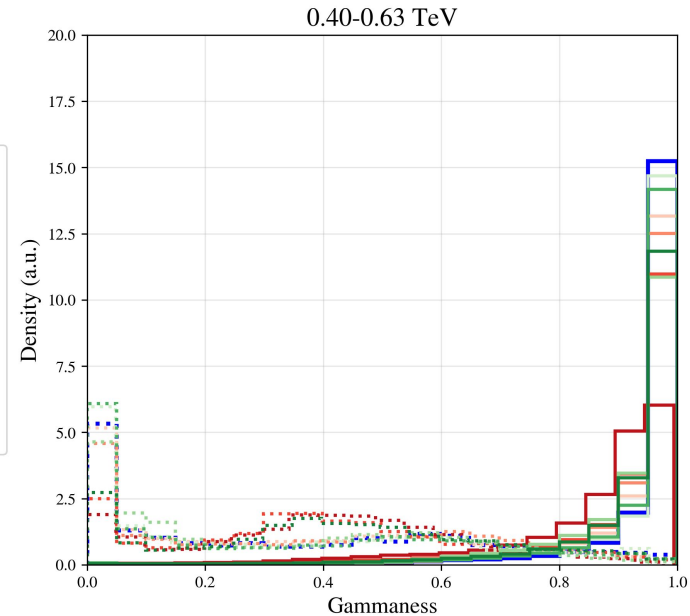
Gammaness Distribution Comparison per Energy Bin



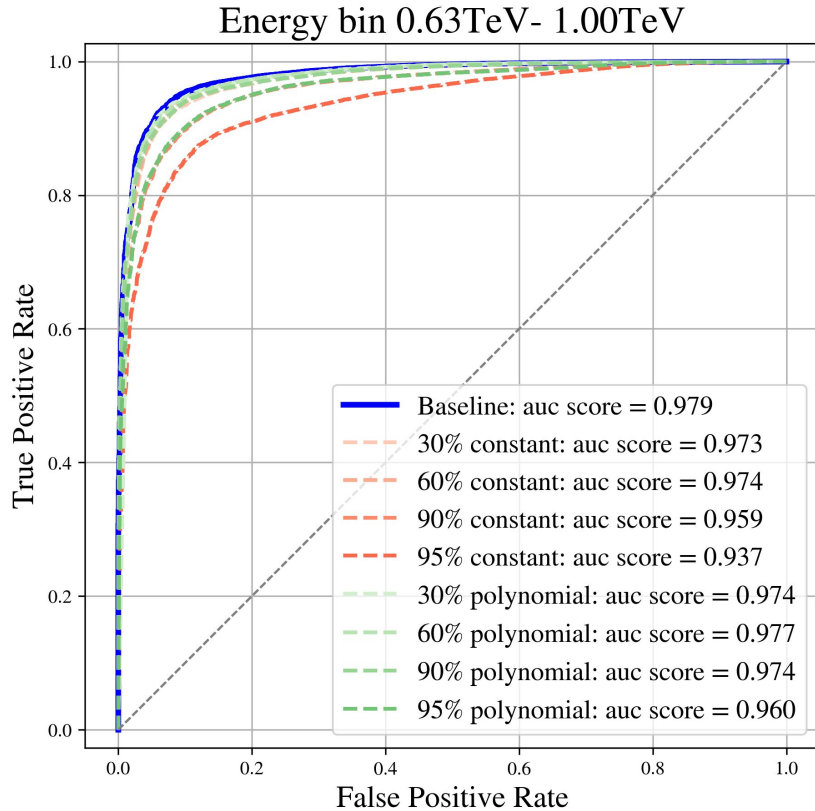
Results: Classification task



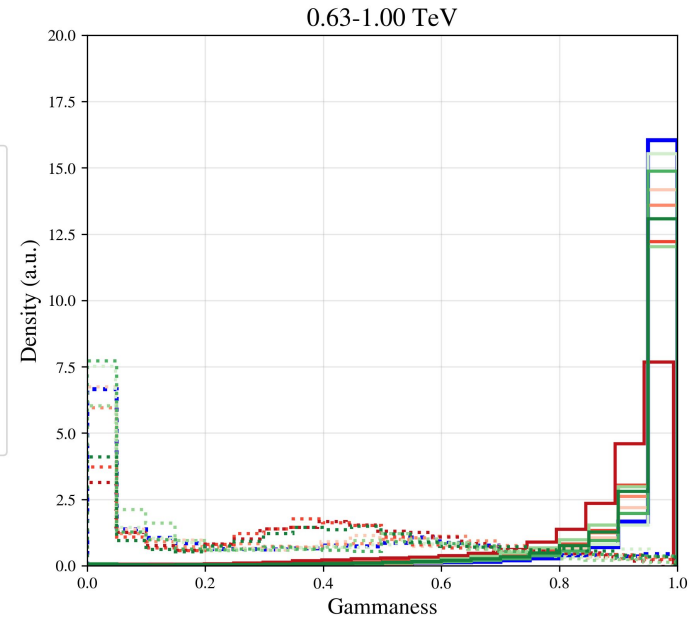
Gammaness Distribution Comparison per Energy Bin



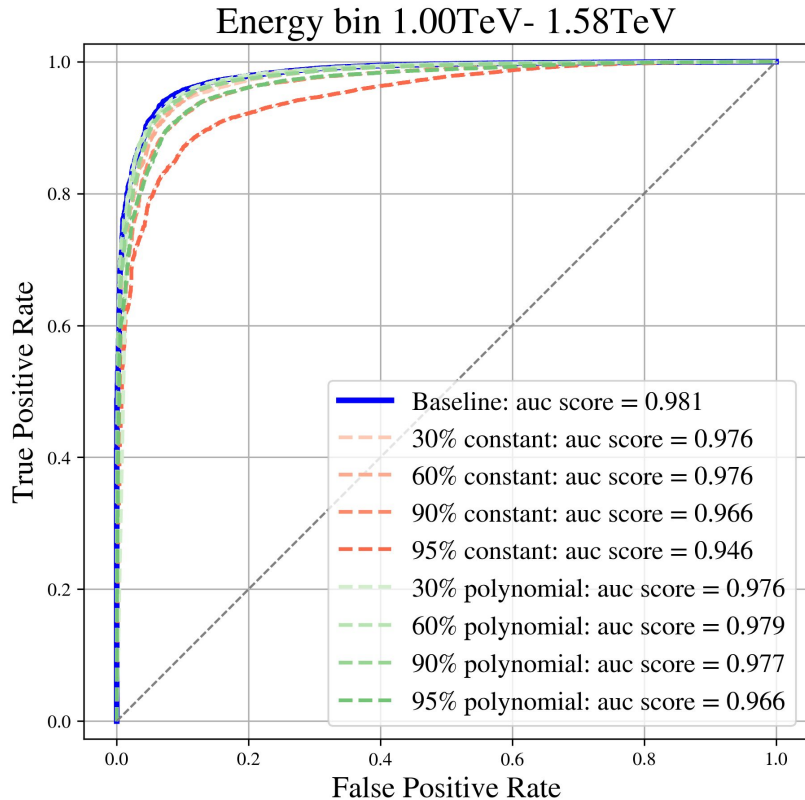
Results: Classification task



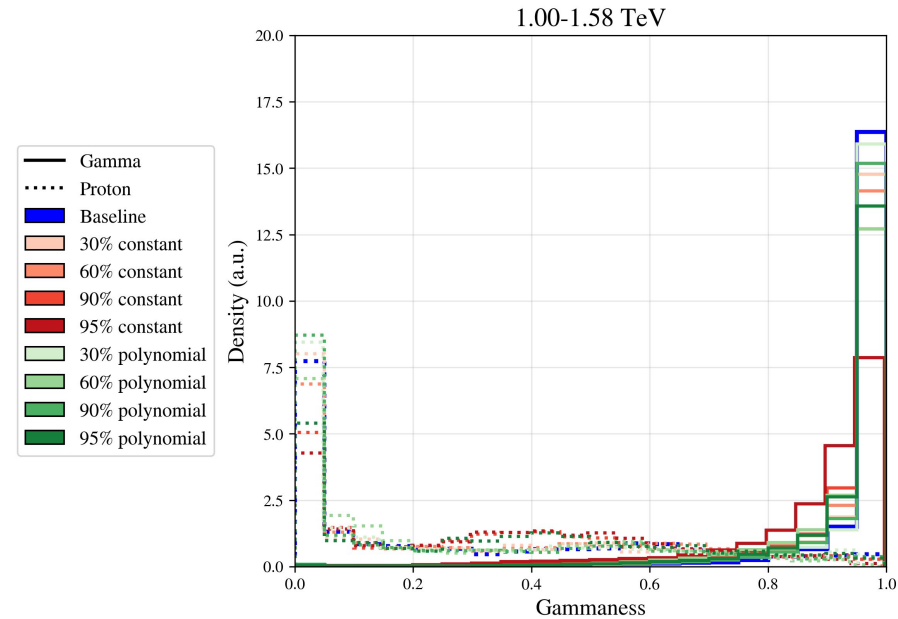
Gammaness Distribution Comparison per Energy Bin



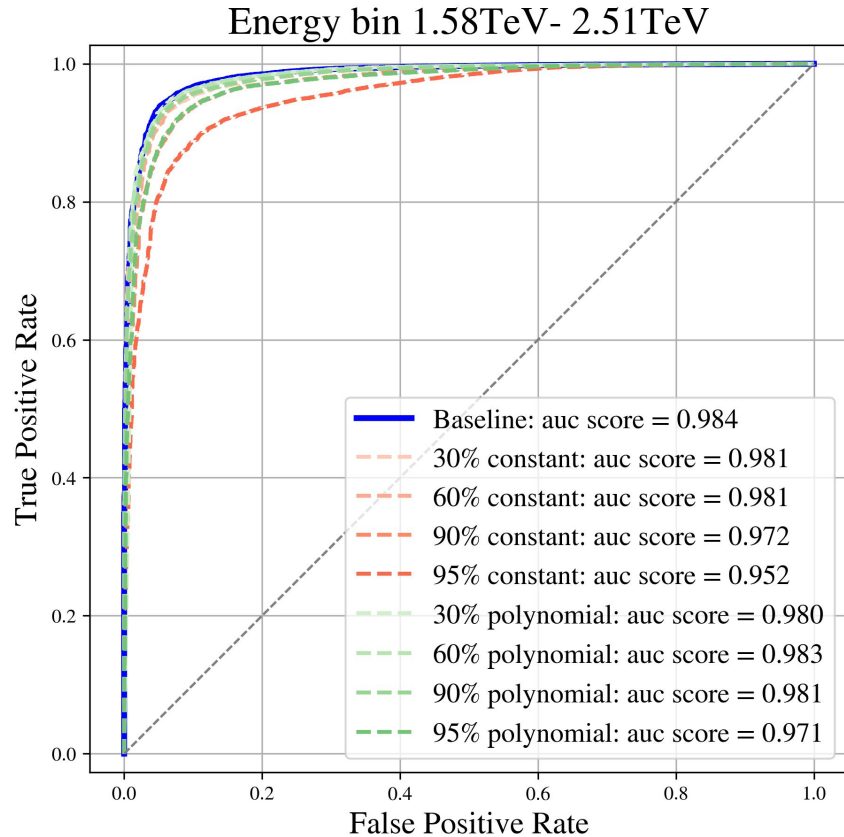
Results: Classification task



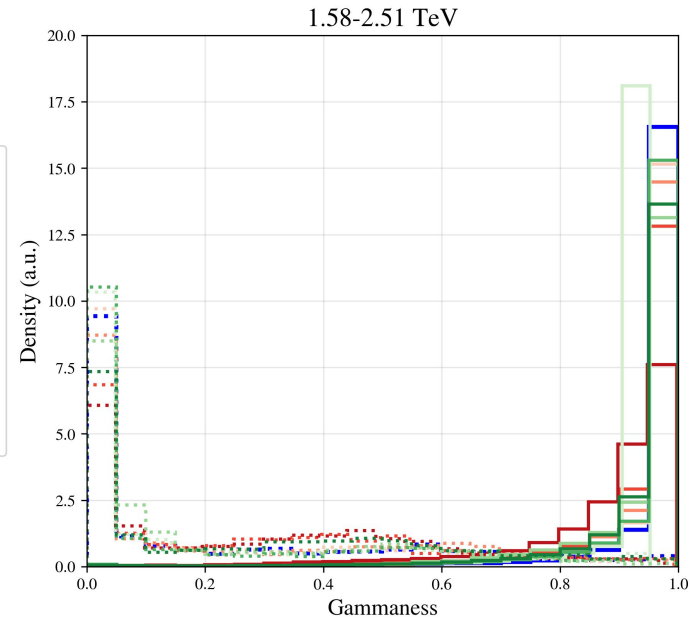
Gammaness Distribution Comparison per Energy Bin



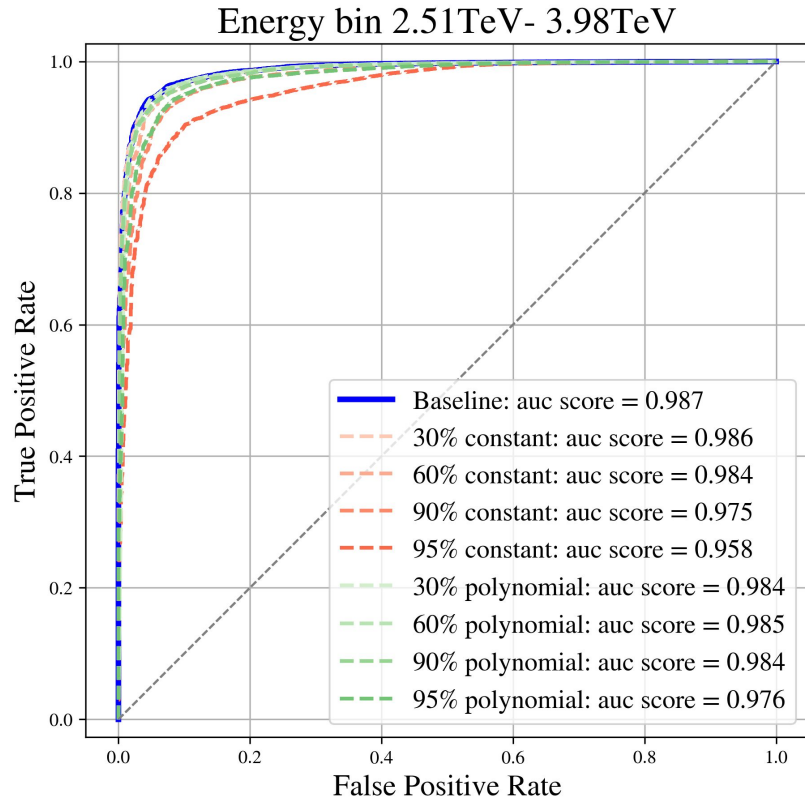
Results: Classification task



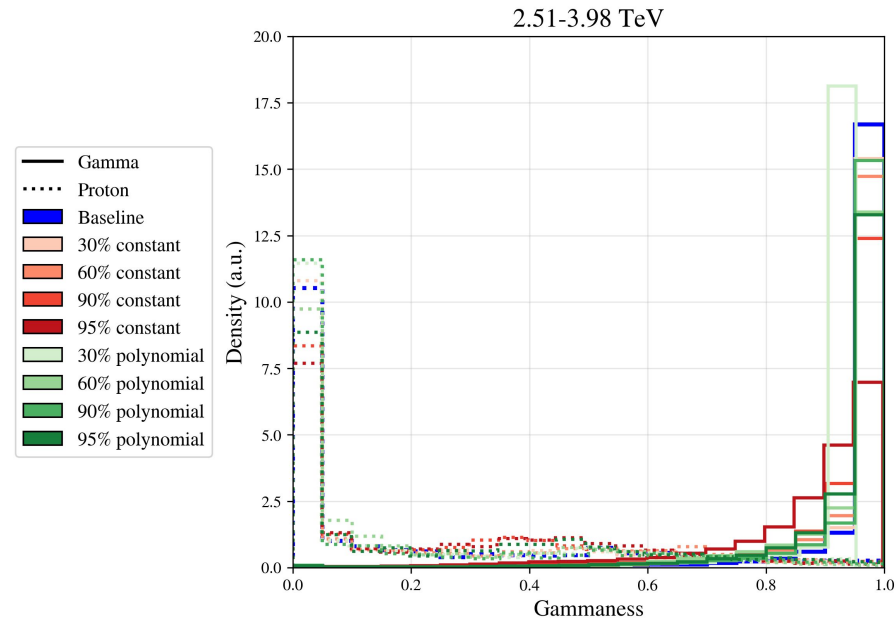
Gammaness Distribution Comparison per Energy Bin



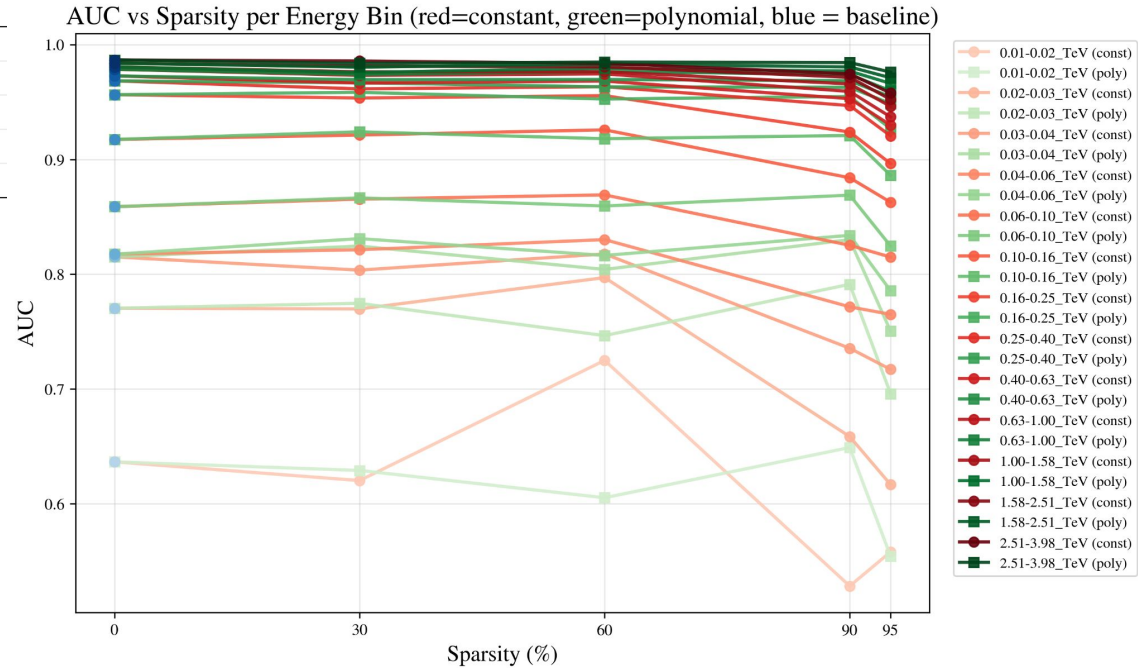
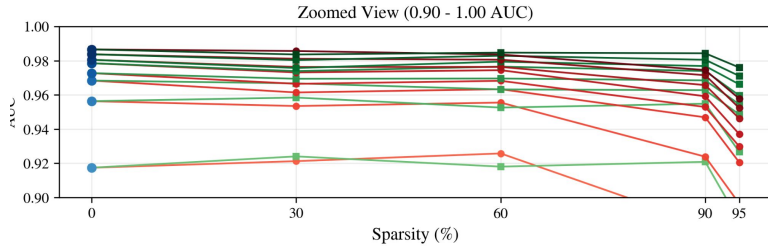
Results: Classification task



Gammaness Distribution Comparison per Energy Bin



Results: Classification task



Results: Regression task

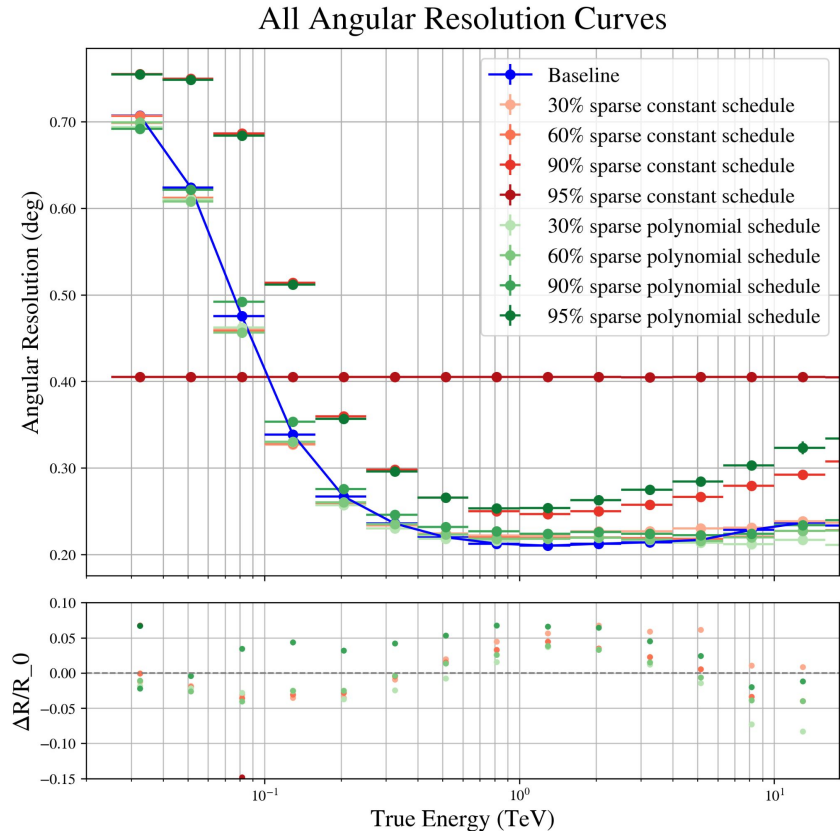


Resolution Curves

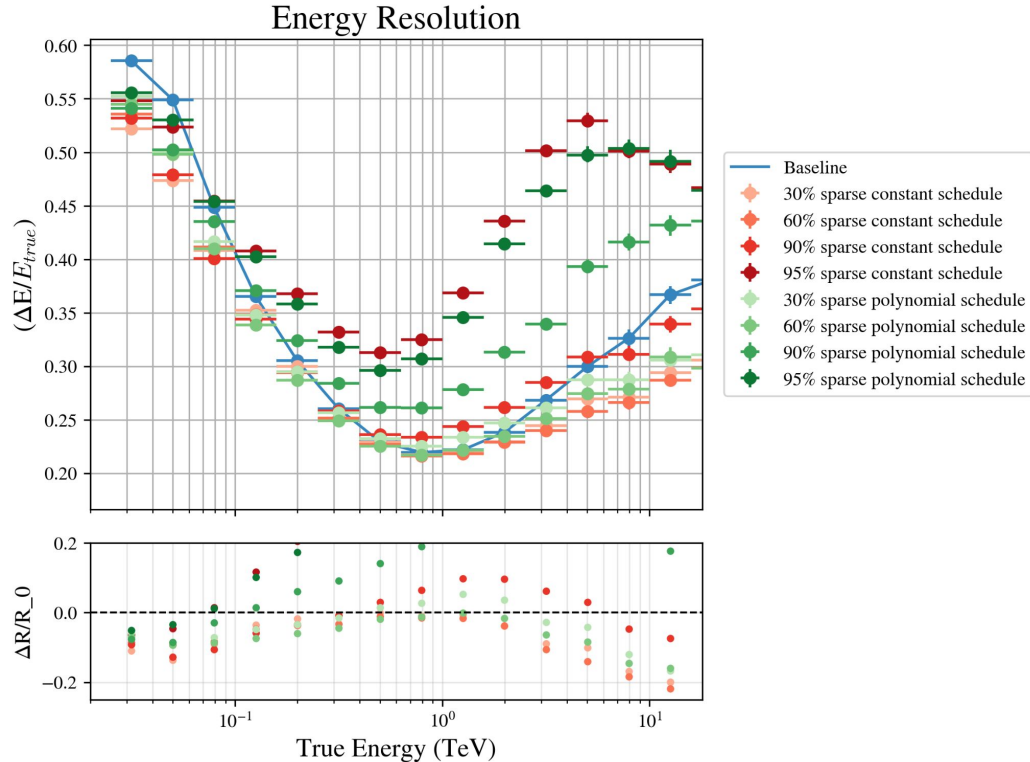
- Angular resolution
- Energy resolution

Results: Regression task

- Polynomial: Similar to baseline until sparsities up to 90%
- Polynomial outperforms constant schedule
- Possible improvement in some energy ranges
- Relative deviation with respect to baseline

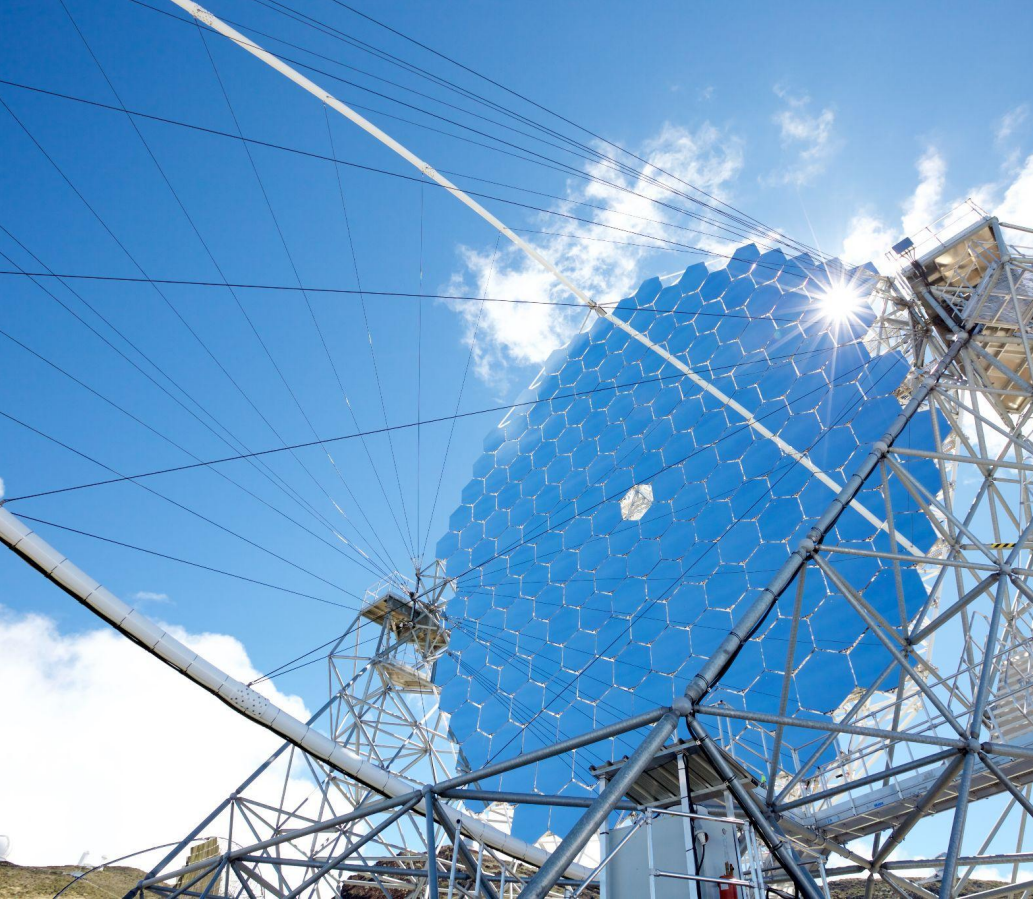


Results: Regression task



- Polynomial: Similar to baseline until sparsities up to 95% de sparsity
- Polynomial outperforms constant schedule
- Possible improvement in some energy ranges
- Relative deviation with respect to baseline

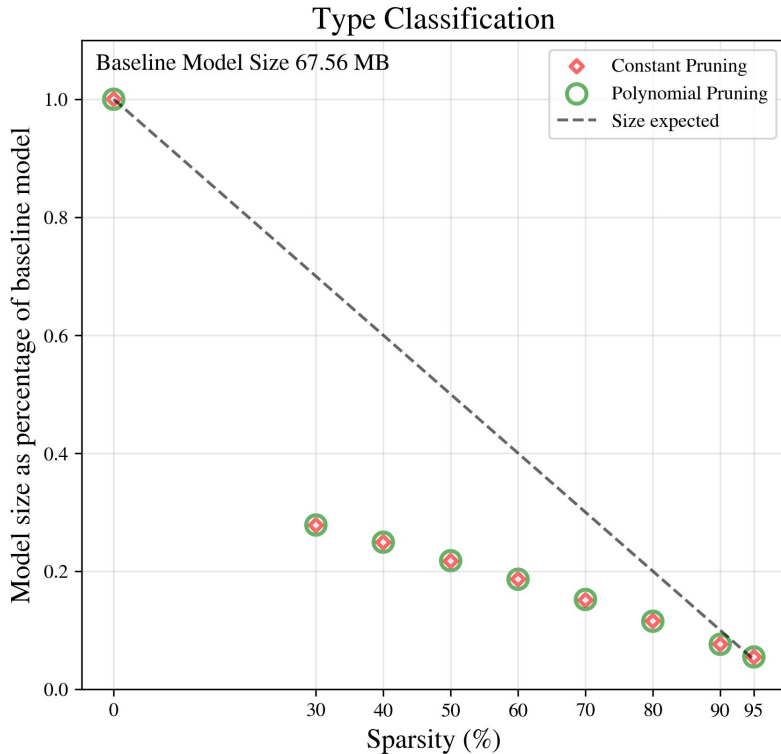
Results: Model size



Evaluate the size of the model as a function of sparsity

Results: Reduction in size

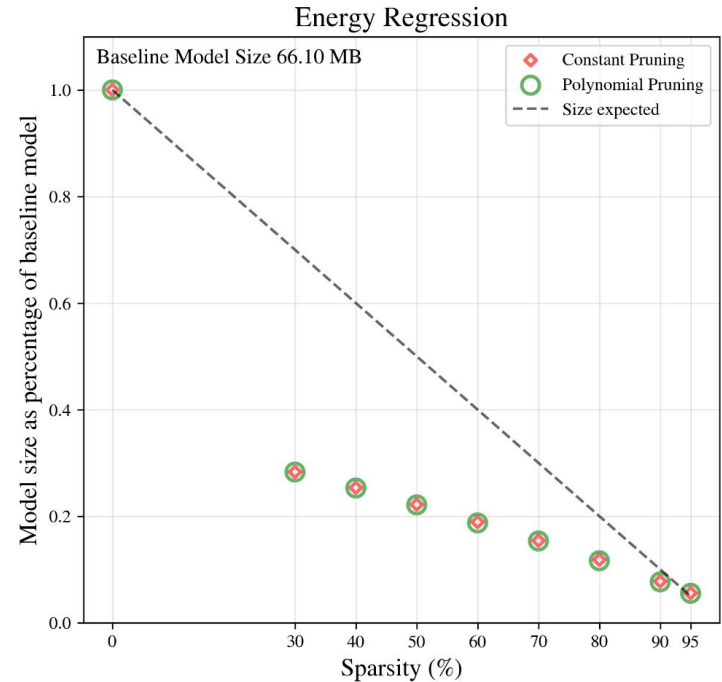
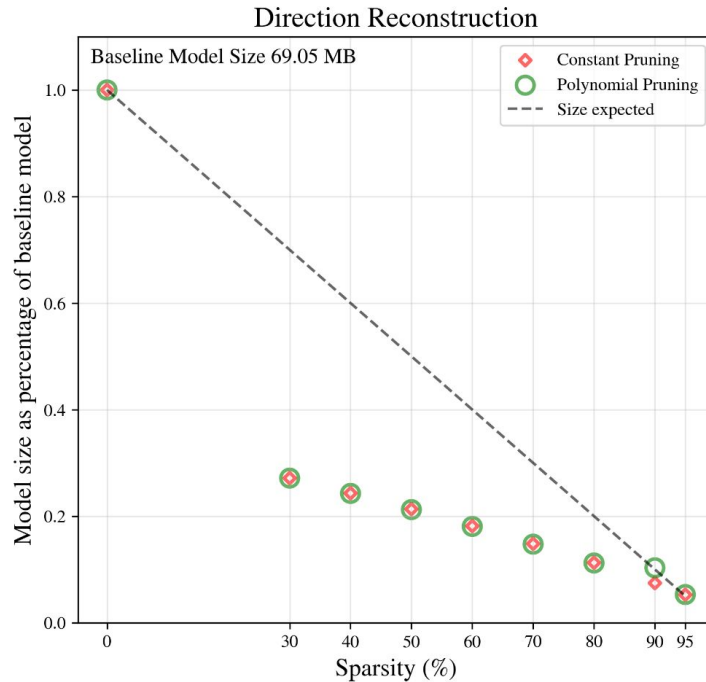
Model Size vs Sparsity for Different Tasks



- Model sizes after compression to .zip
- Smaller size than linear reduction
- 30% sparsity: 28% of original size
- 95% sparsity: 5% of original size

Results: Reduction in size

Model Size vs Sparsity for Different Tasks



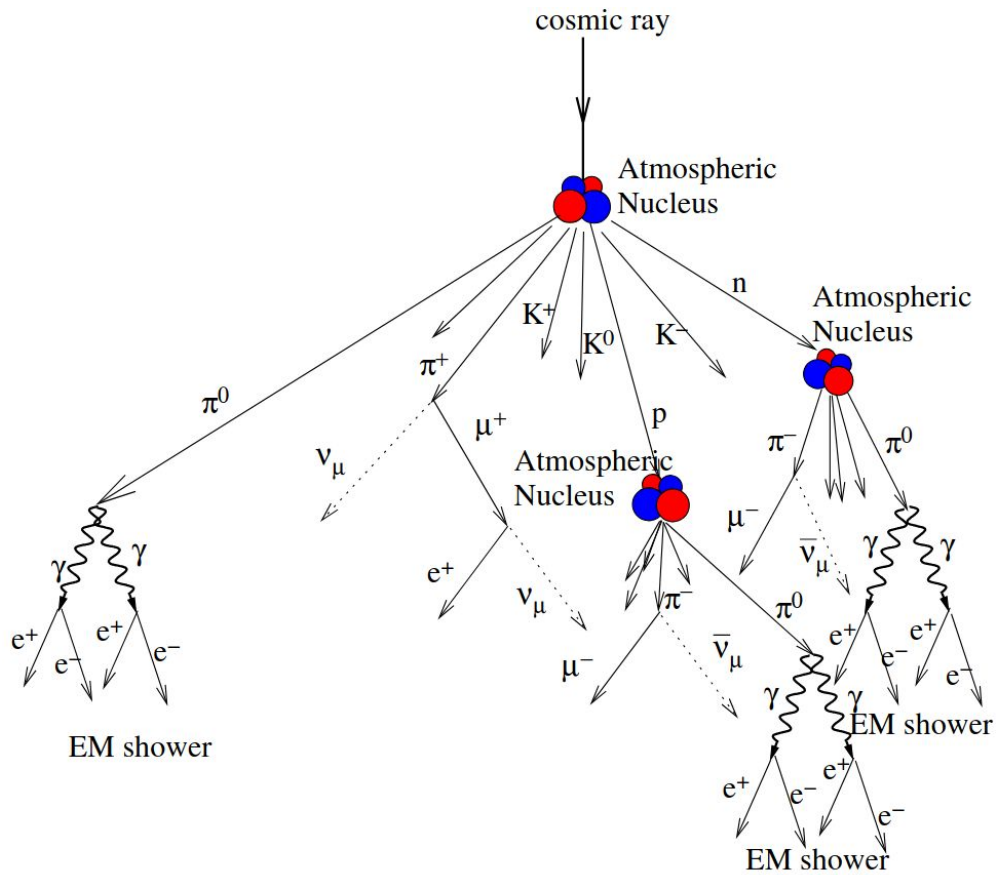
Conclusions & Future Work

Conclusions

- **Polynomial sparsity** outperforms **constant sparsity**
- **Classification: polynomial schedule** without significant degradation up to 90-95%.
- **Energy polynomial schedule** without significant degradation at 90%.
- **Direction: polynomial schedule** without significant degradation up to 90%.
- Significant reduction in model **size**

Ongoing Work

- NSB vs. cosmics classification
- L1 over flowers with 2155 threshold
- Use simpler CNN model
- Sparsity-aware quantisation implementation

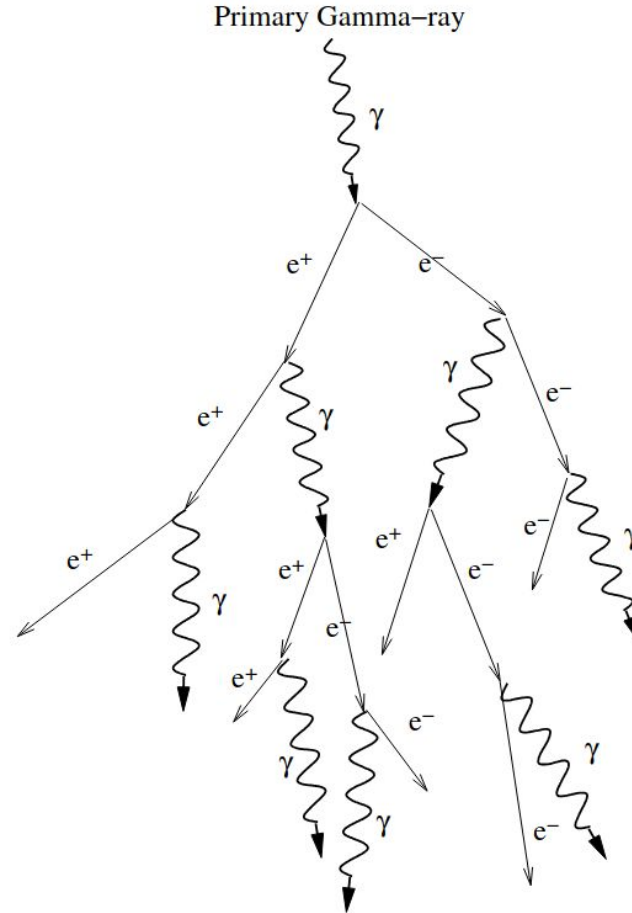


Cascada Hadrónica:

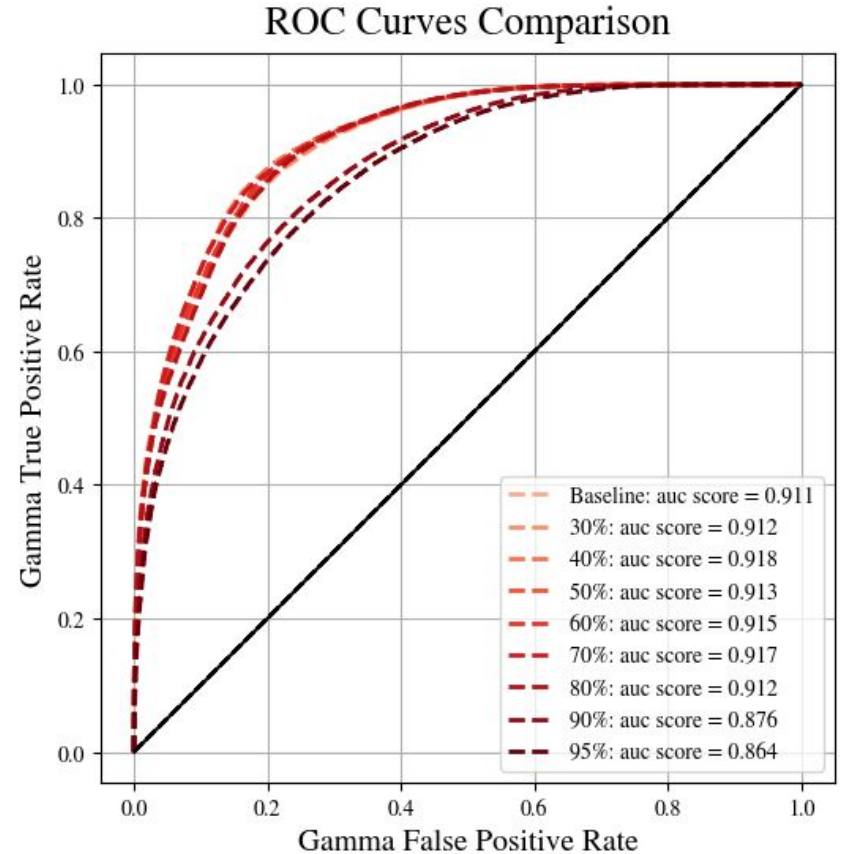
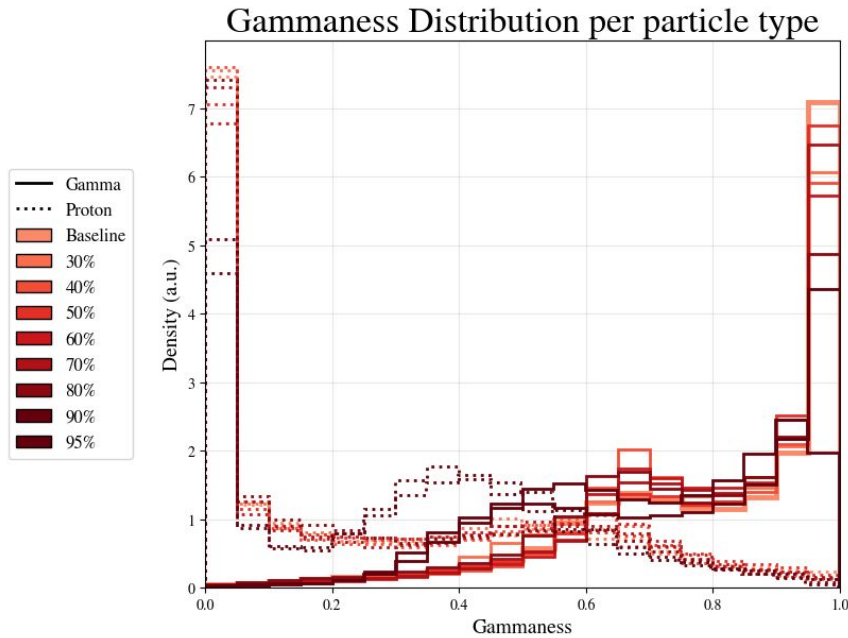
- Protón
- Núcleos de Helio (α)

Cascada Electromagnética:

- Rayo Gamma (γ)

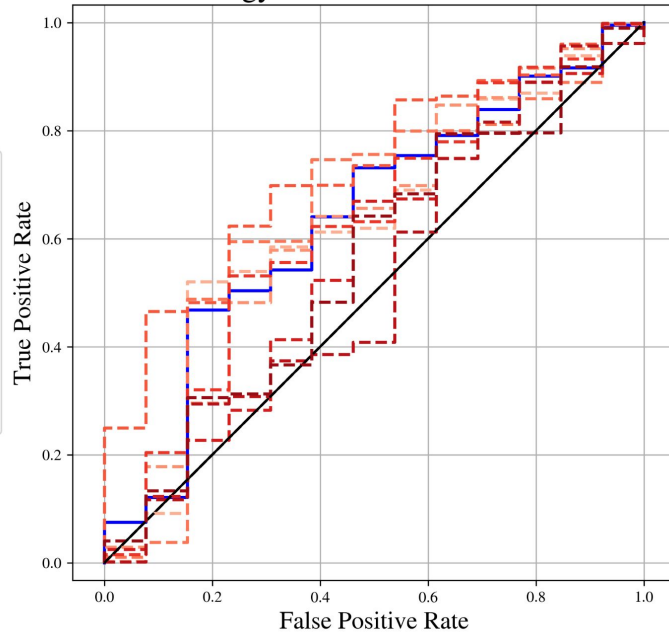


Resultados: Tarea de clasificación

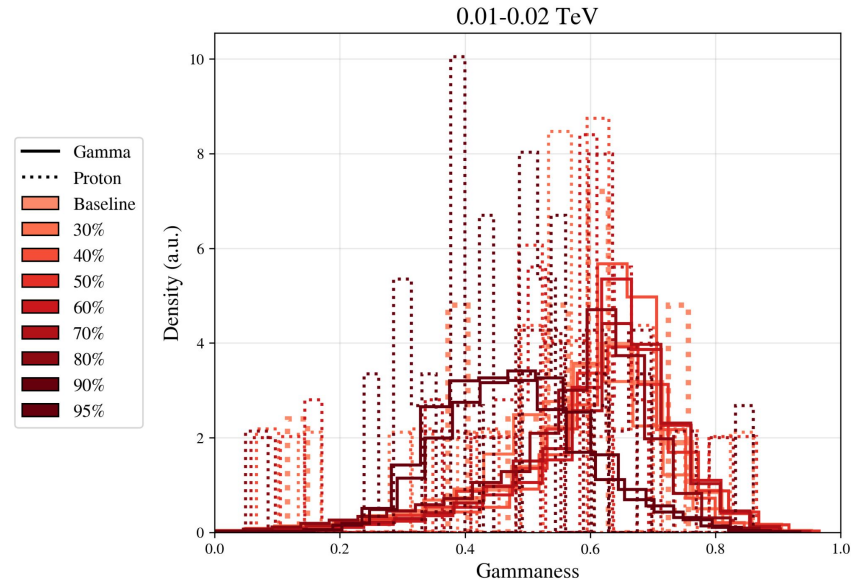


Resultados: Tarea de clasificación

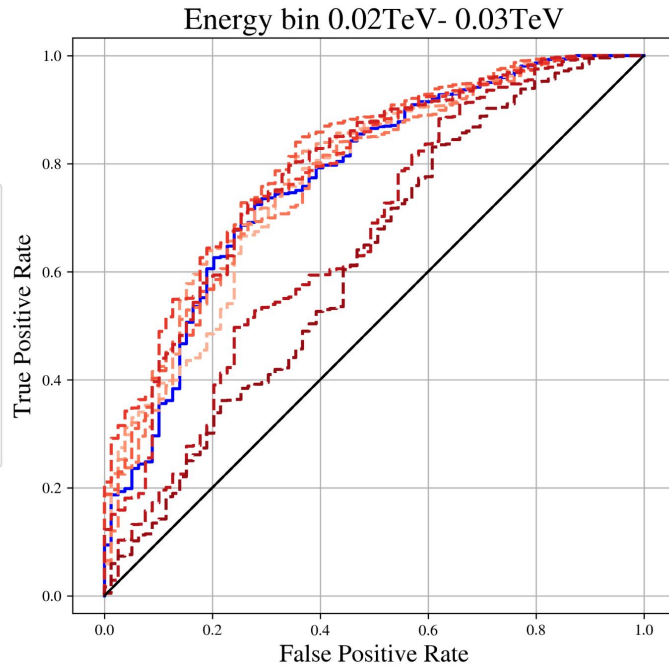
Energy bin 0.01TeV- 0.02TeV



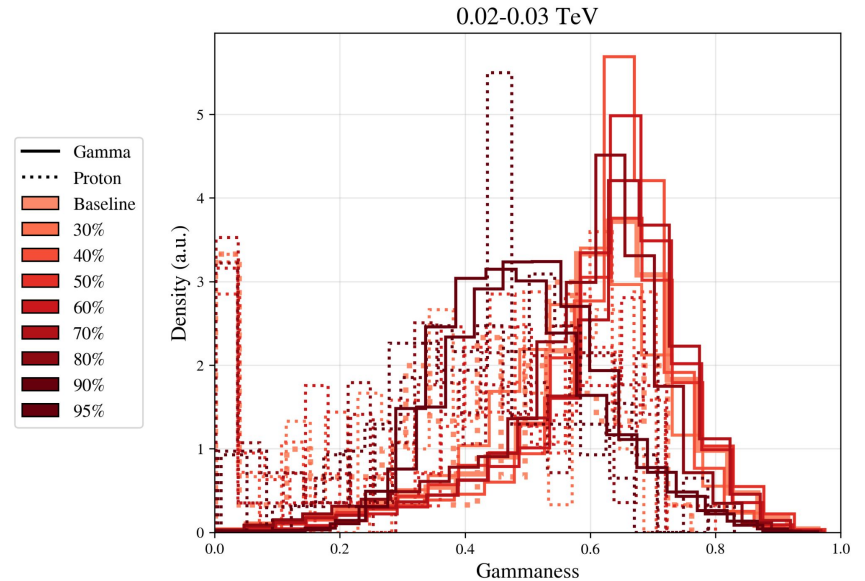
Gammaness Distribution Comparison per Energy Bin



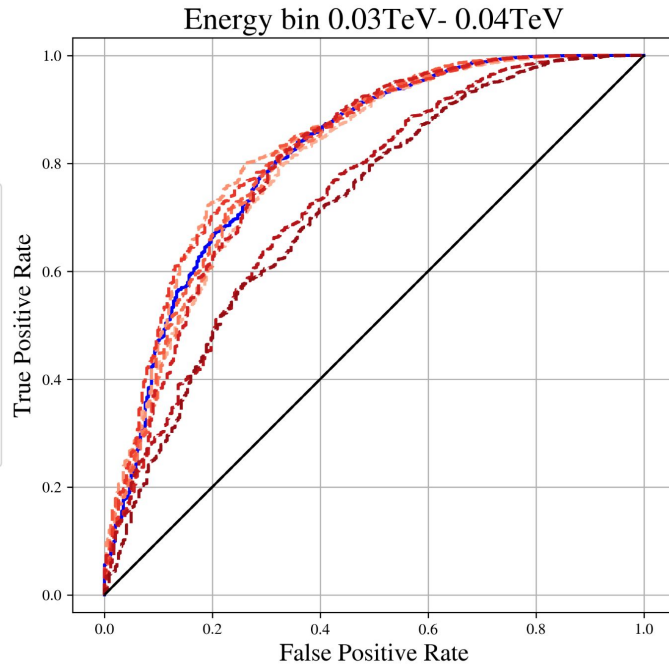
Resultados: Tarea de clasificación



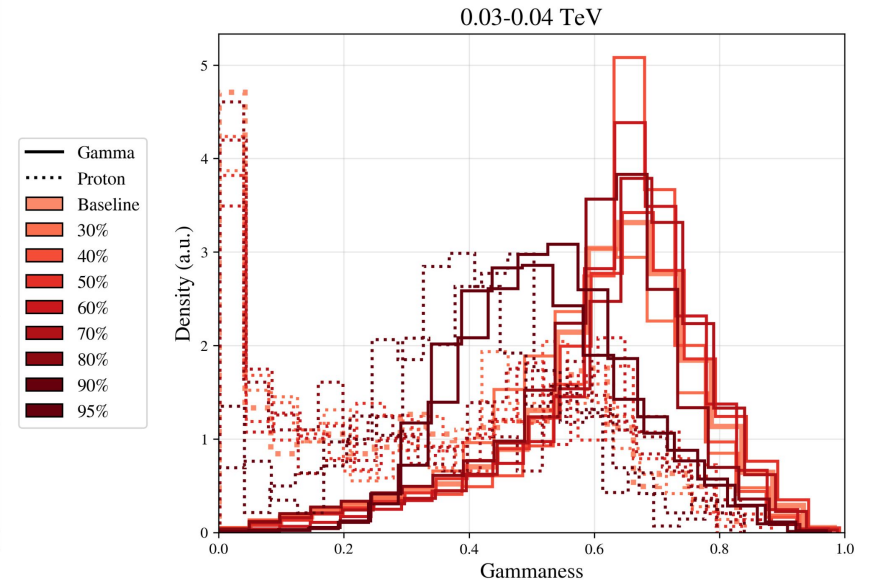
Gammaness Distribution Comparison per Energy Bin



Resultados: Tarea de clasificación

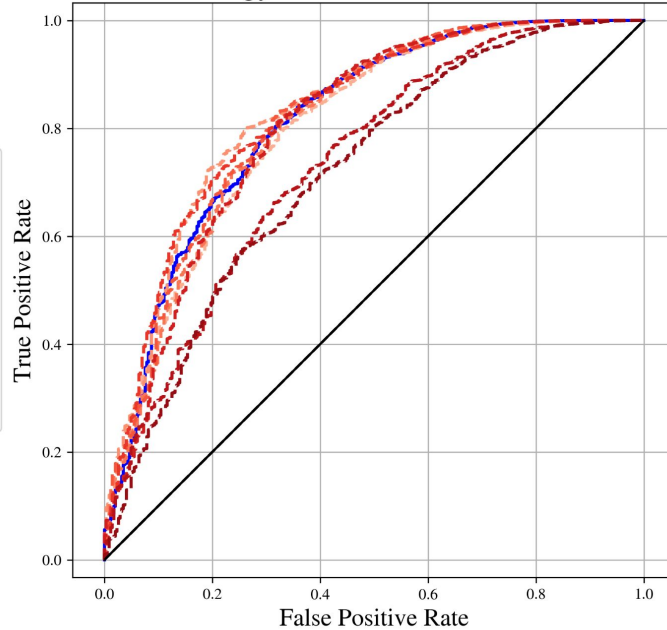


Gammaness Distribution Comparison per Energy Bin

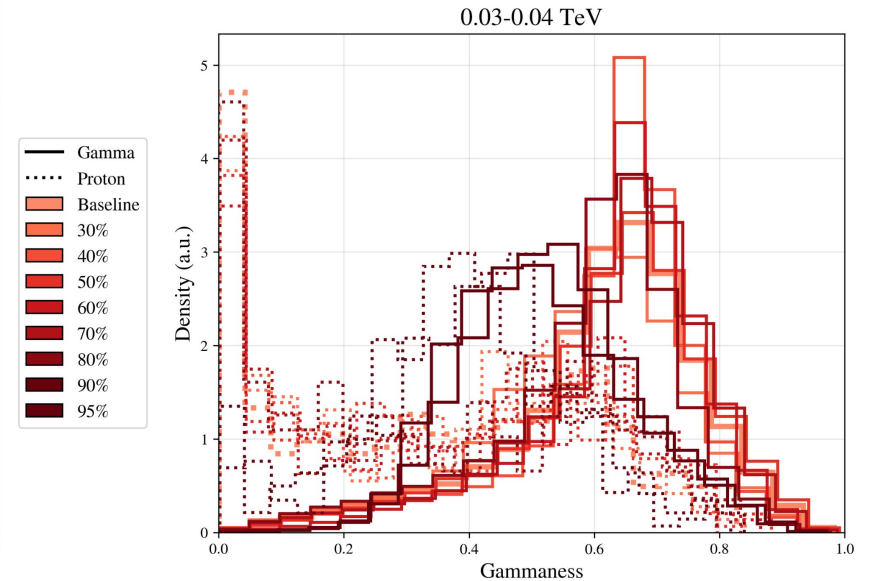


Resultados: Tarea de clasificación

Energy bin 0.03TeV- 0.04TeV

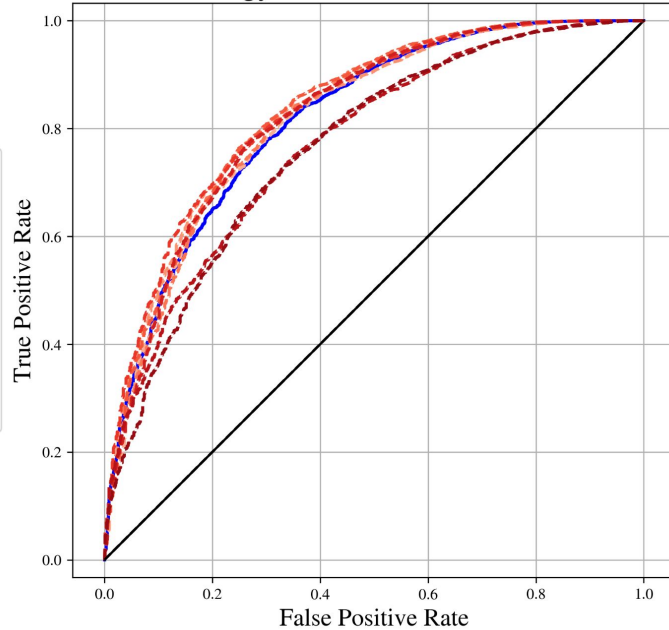


Gammaness Distribution Comparison per Energy Bin

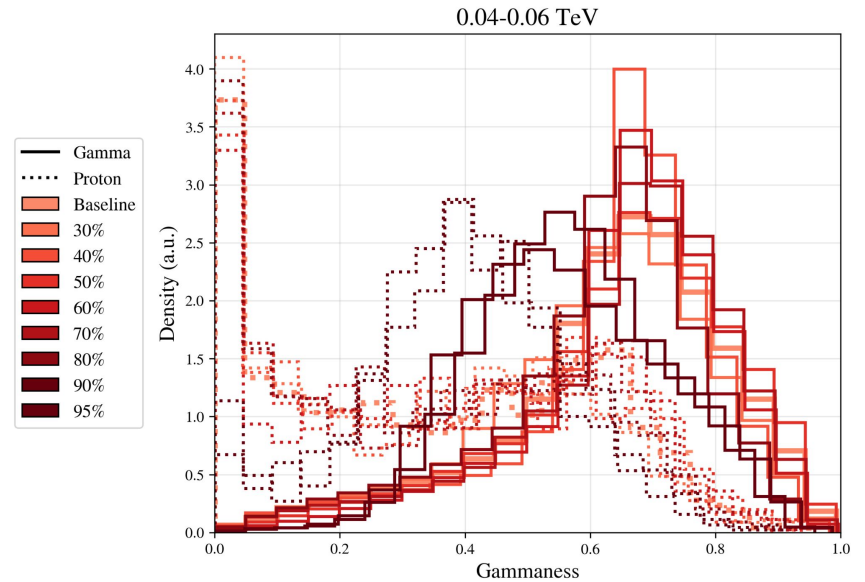


Resultados: Tarea de clasificación

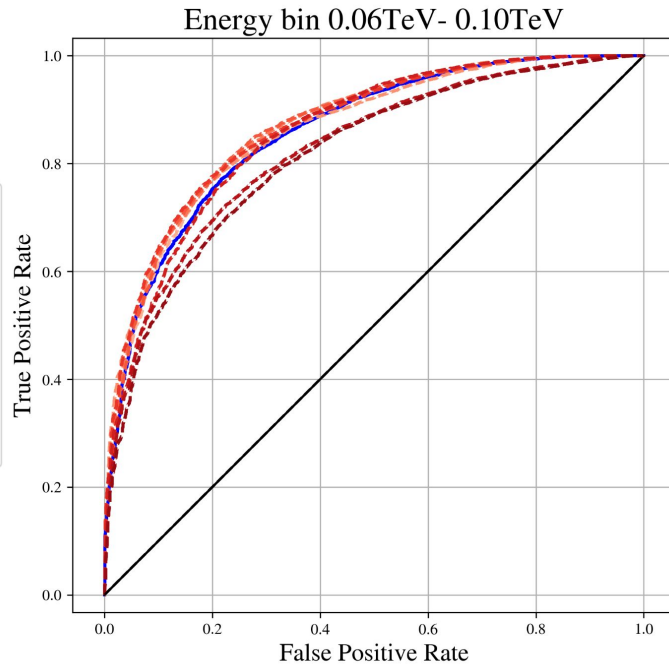
Energy bin 0.04TeV- 0.06TeV



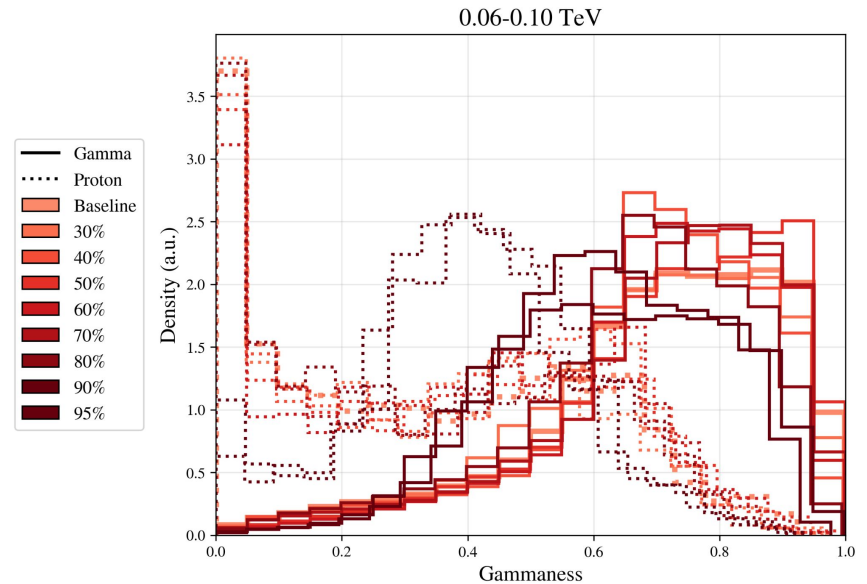
Gammaness Distribution Comparison per Energy Bin



Resultados: Tarea de clasificación

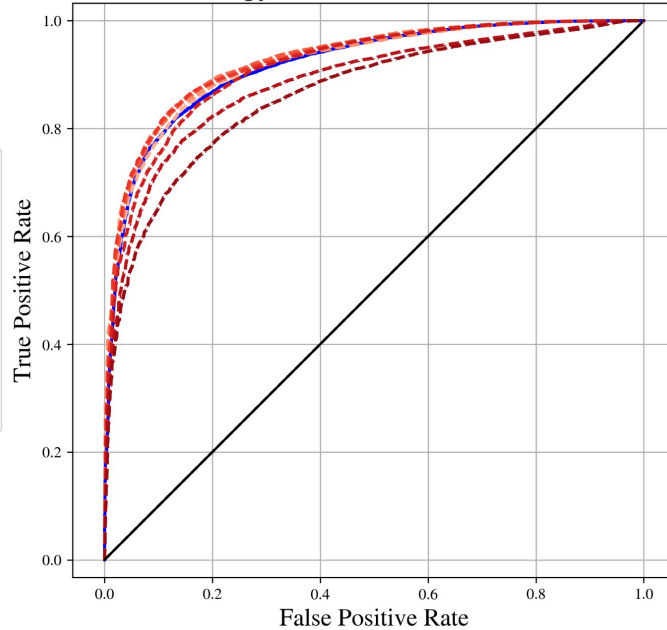


Gammaness Distribution Comparison per Energy Bin



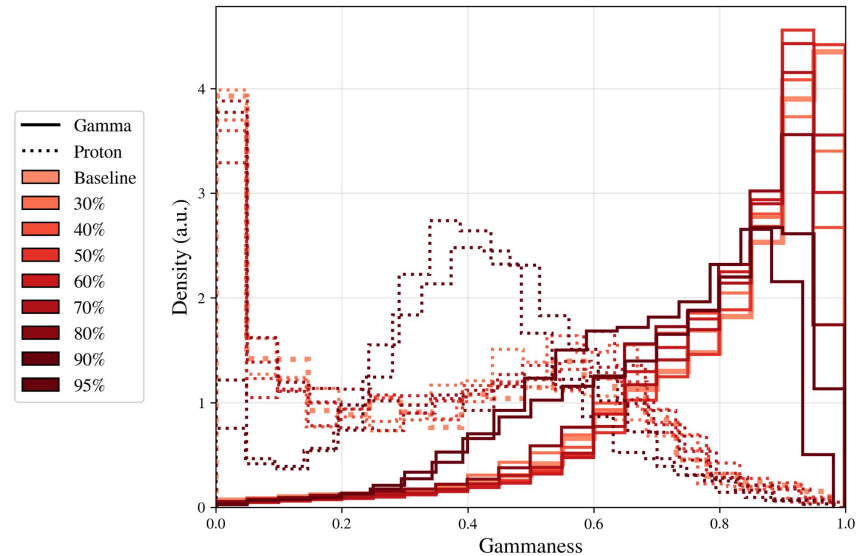
Resultados: Tarea de clasificación

Energy bin 0.10TeV- 0.16TeV

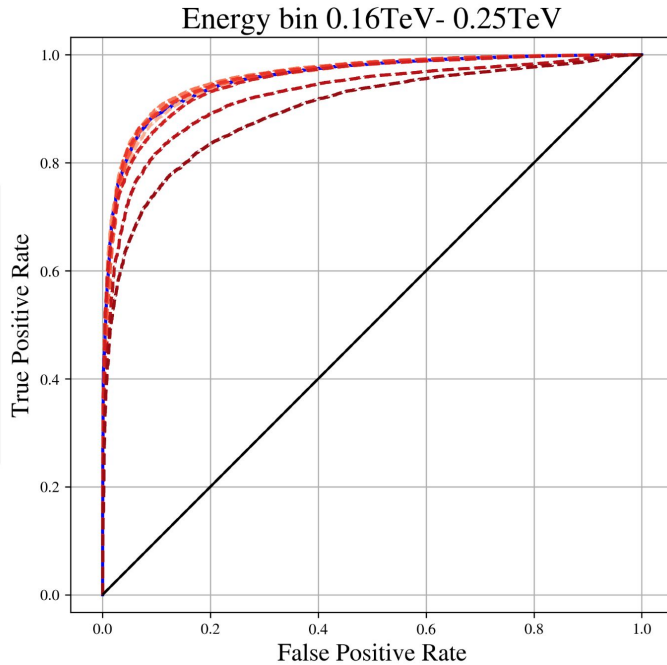


Gammaness Distribution Comparison per Energy Bin

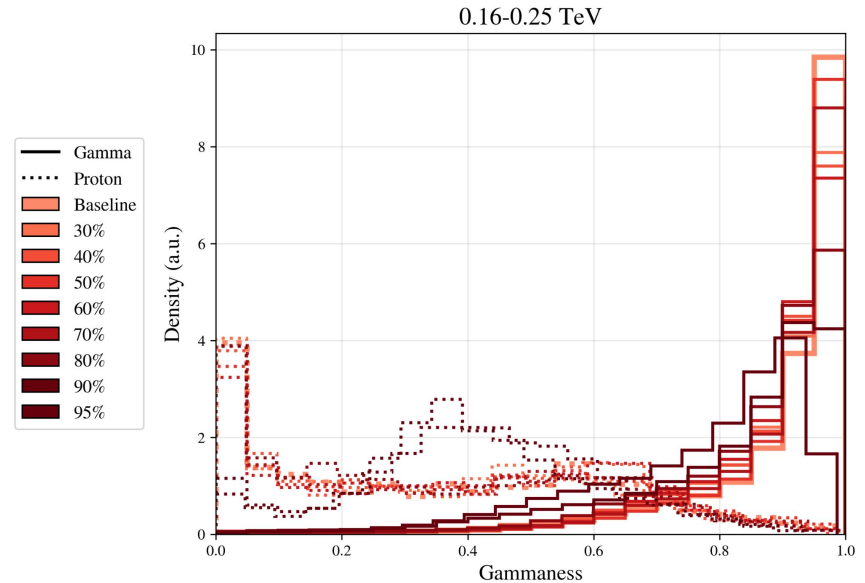
0.10-0.16 TeV



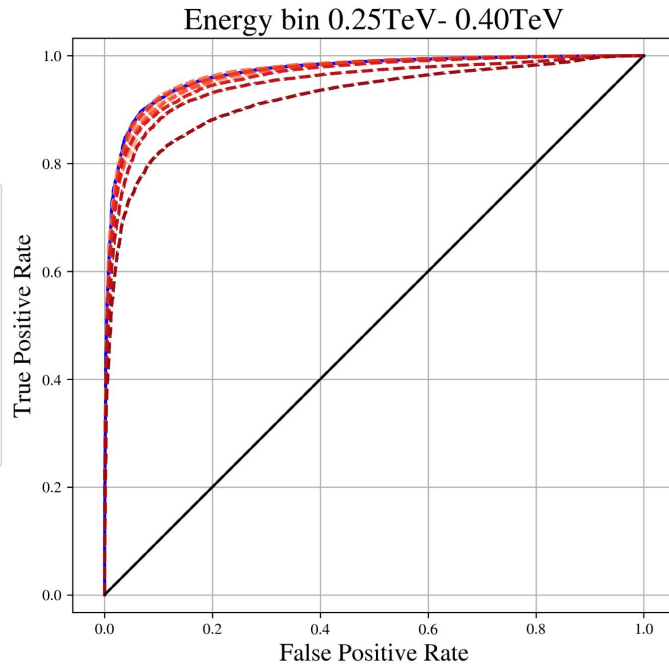
Resultados: Tarea de clasificación



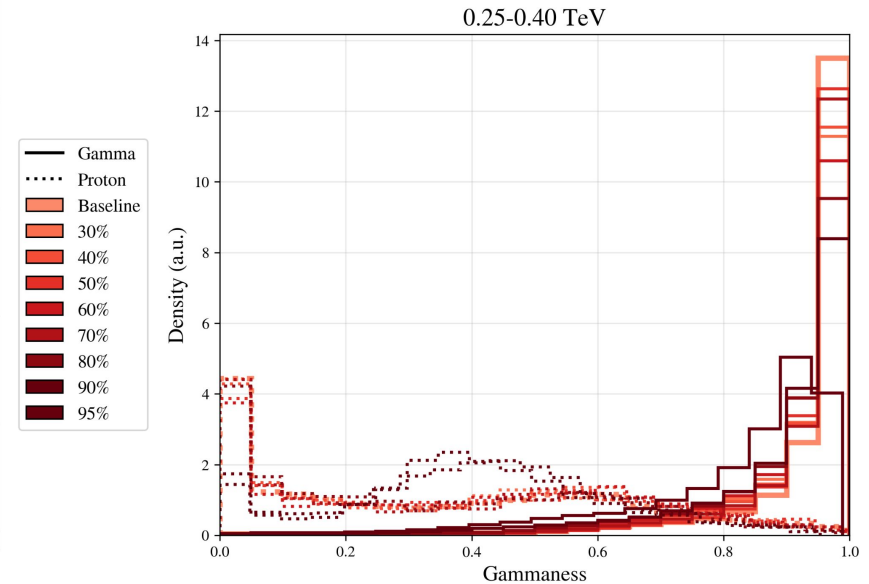
Gammaness Distribution Comparison per Energy Bin



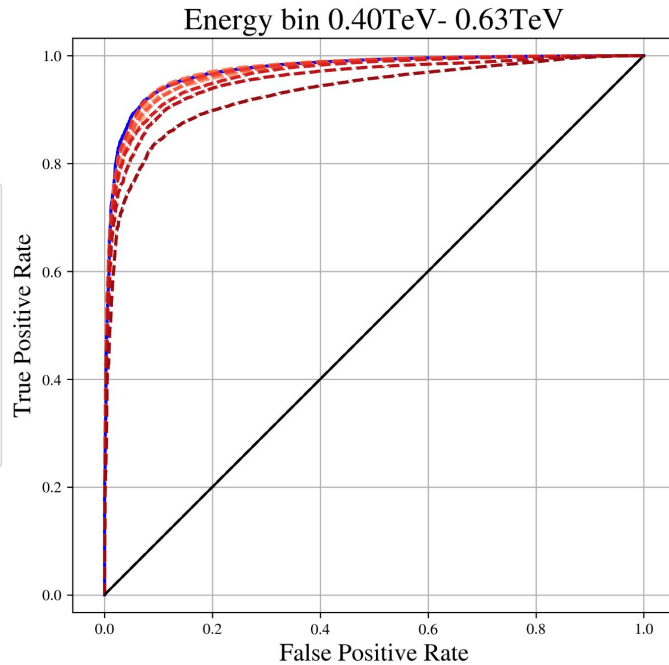
Resultados: Tarea de clasificación



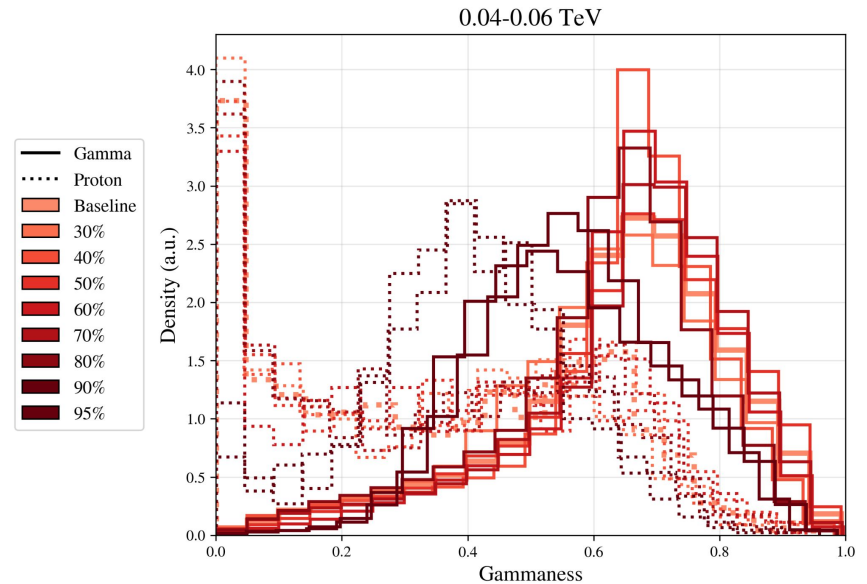
Gammaness Distribution Comparison per Energy Bin



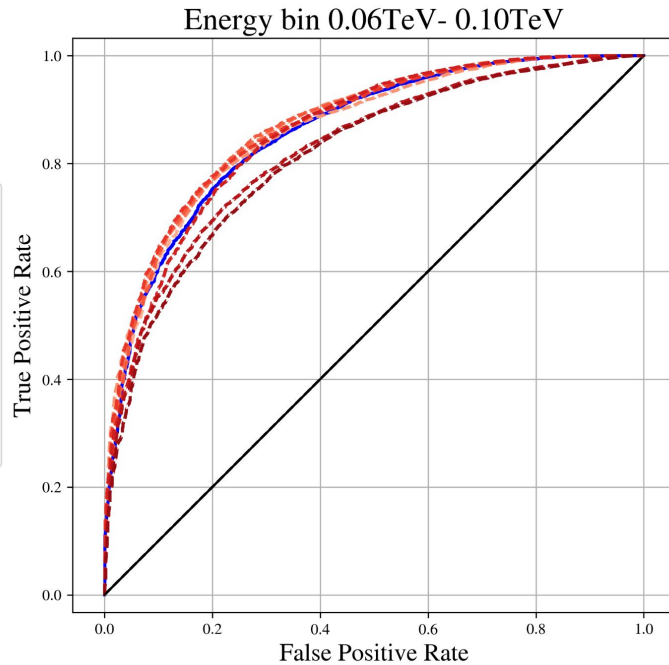
Resultados: Tarea de clasificación



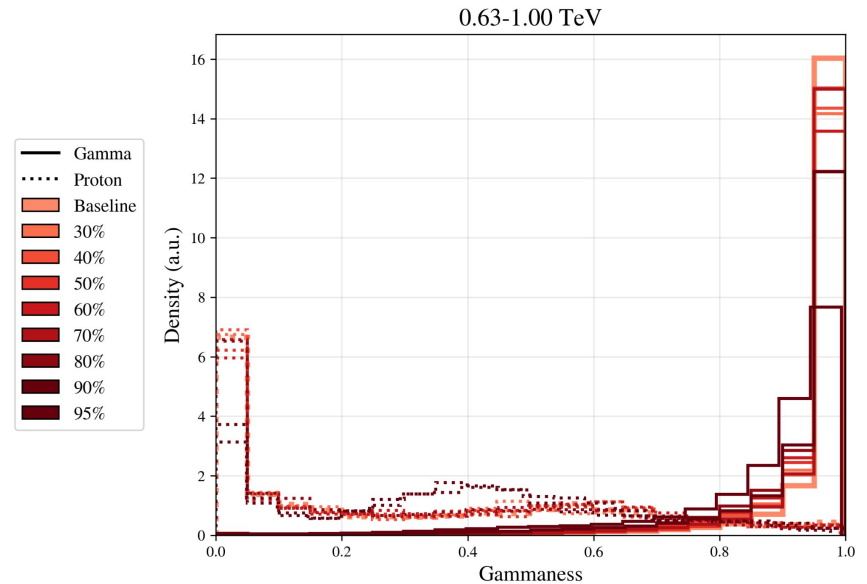
Gammaness Distribution Comparison per Energy Bin



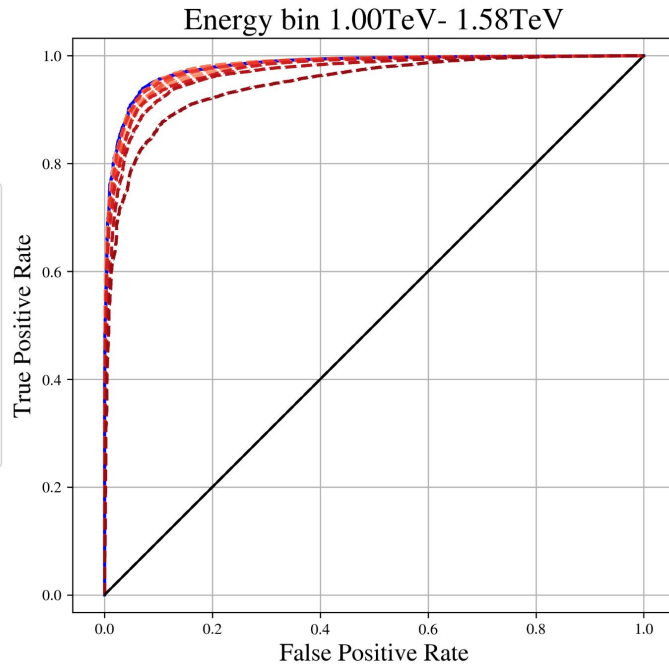
Resultados: Tarea de clasificación



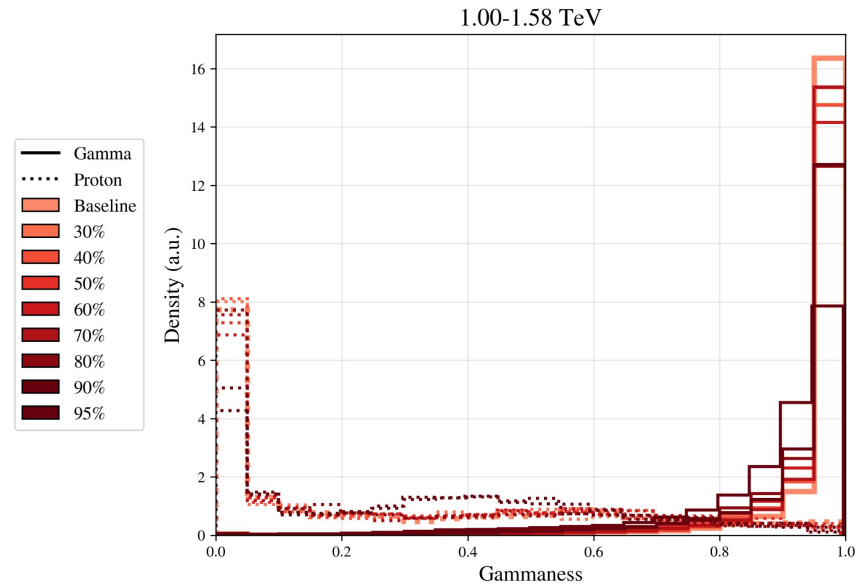
Gammaness Distribution Comparison per Energy Bin



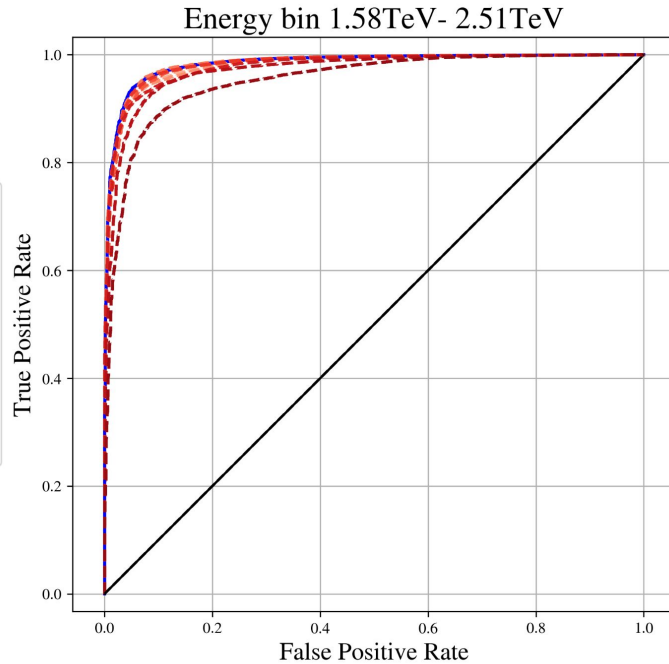
Resultados: Tarea de clasificación



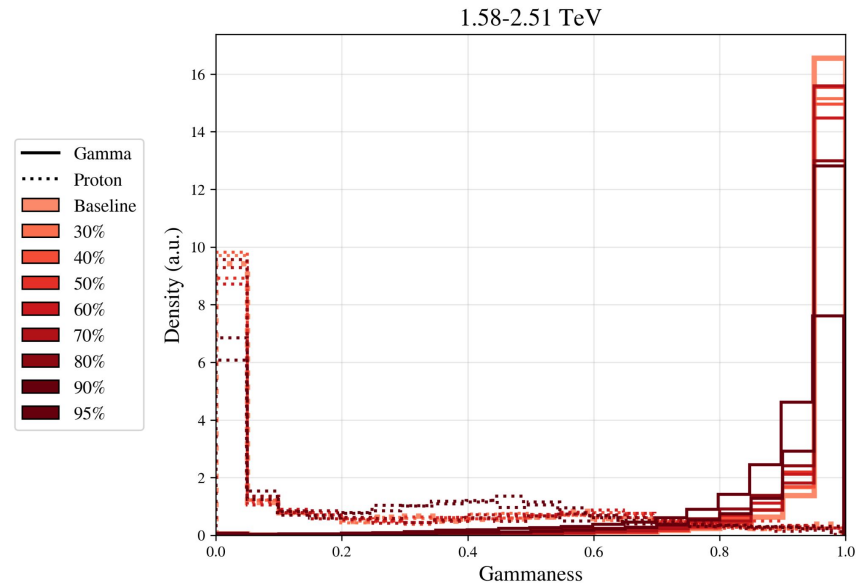
Gammaness Distribution Comparison per Energy Bin



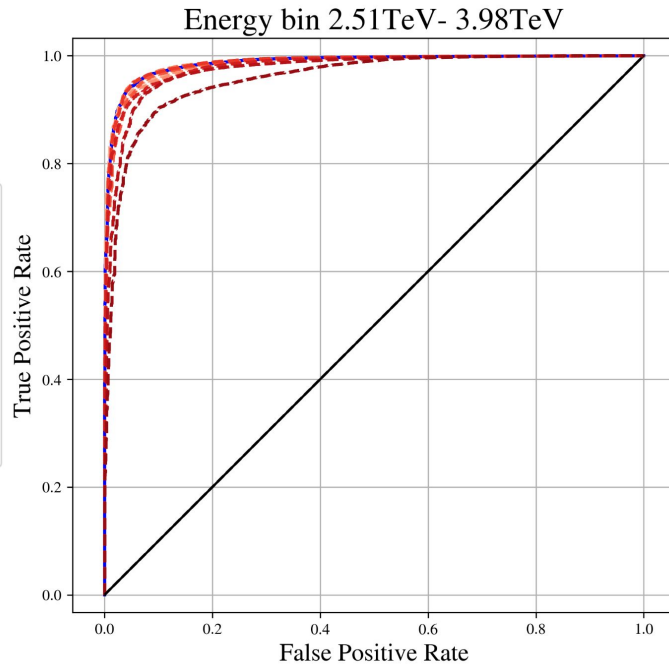
Resultados: Tarea de clasificación



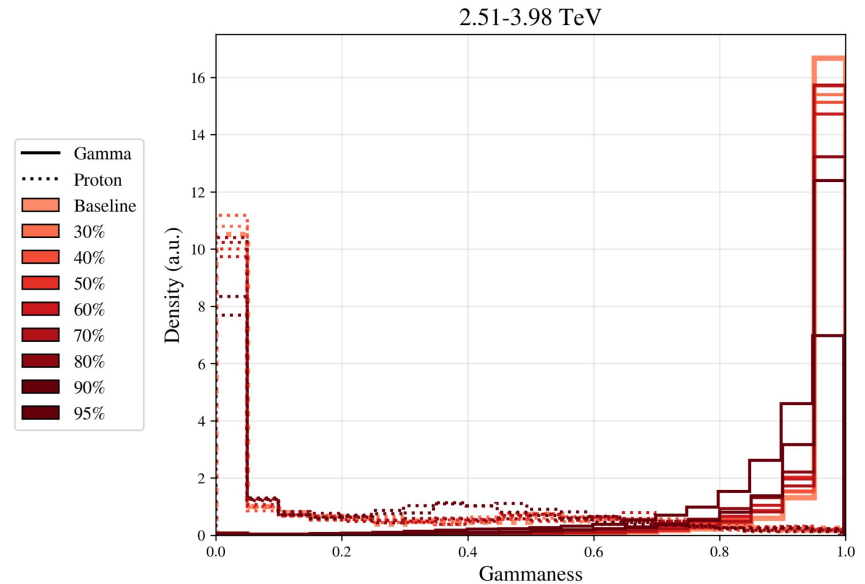
Gammaness Distribution Comparison per Energy Bin



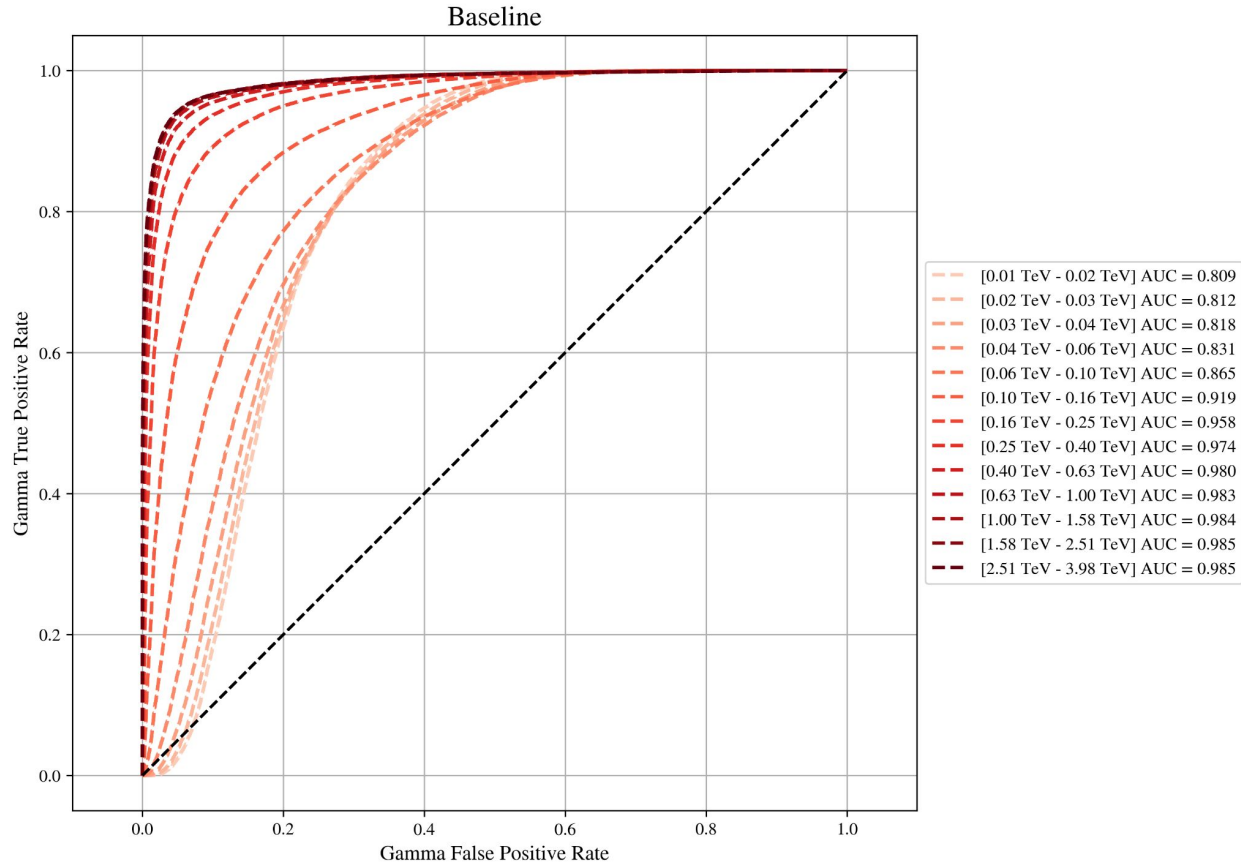
Resultados: Tarea de clasificación



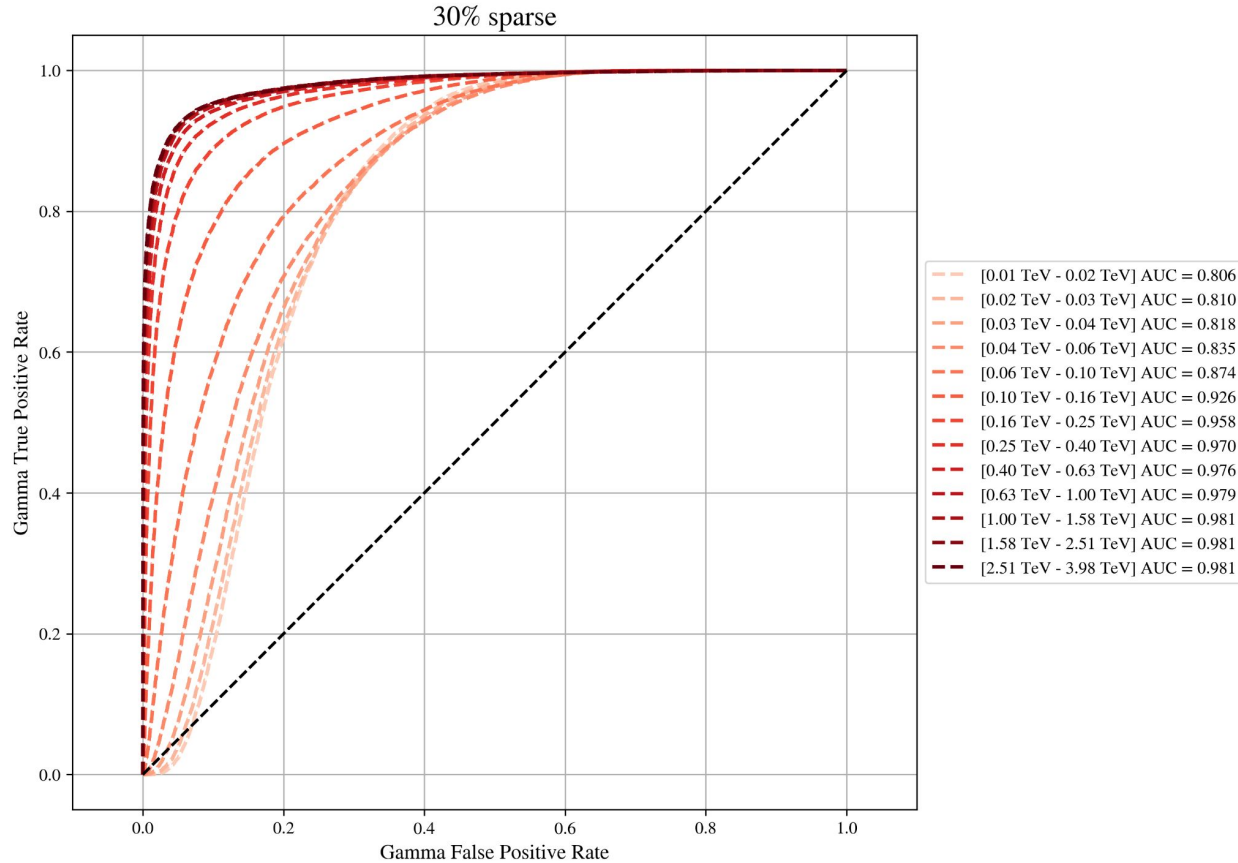
Gammaness Distribution Comparison per Energy Bin



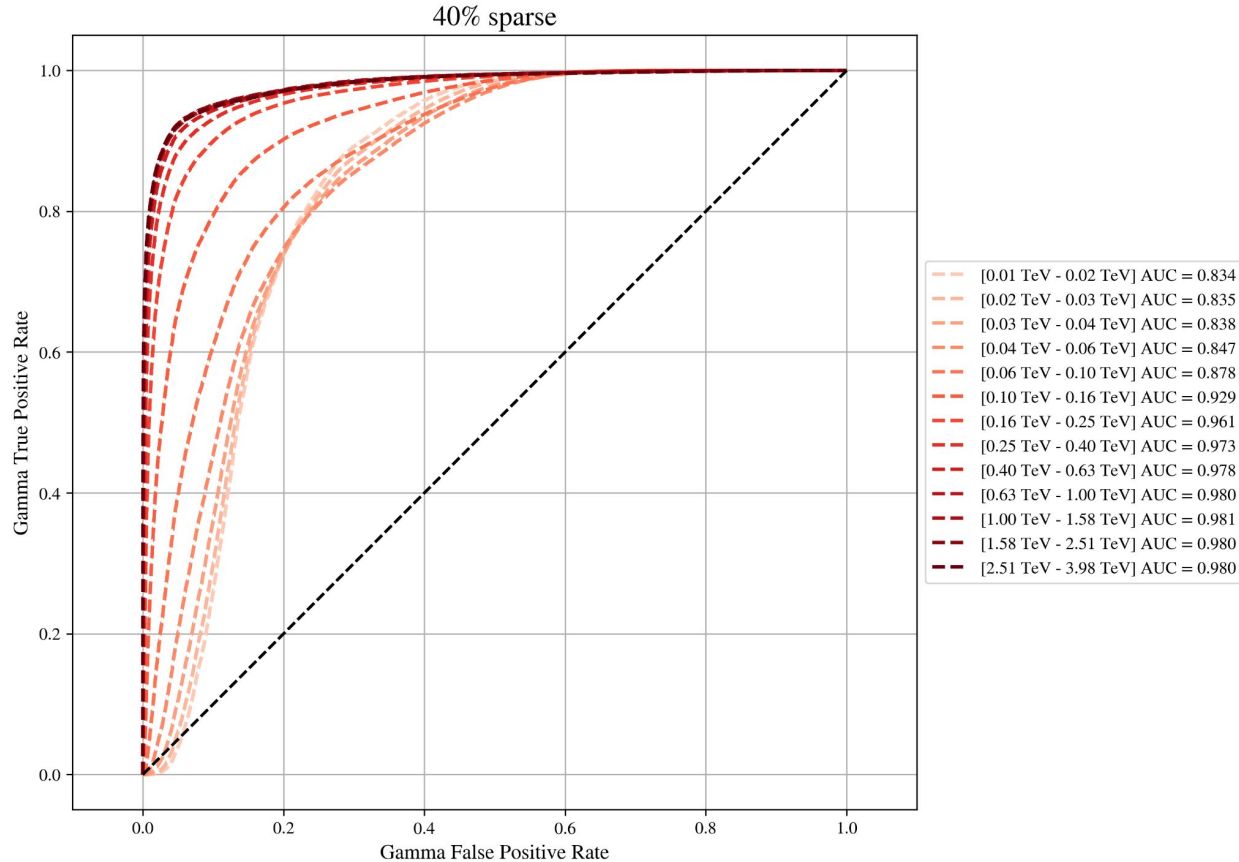
Resultados: Tarea de clasificación



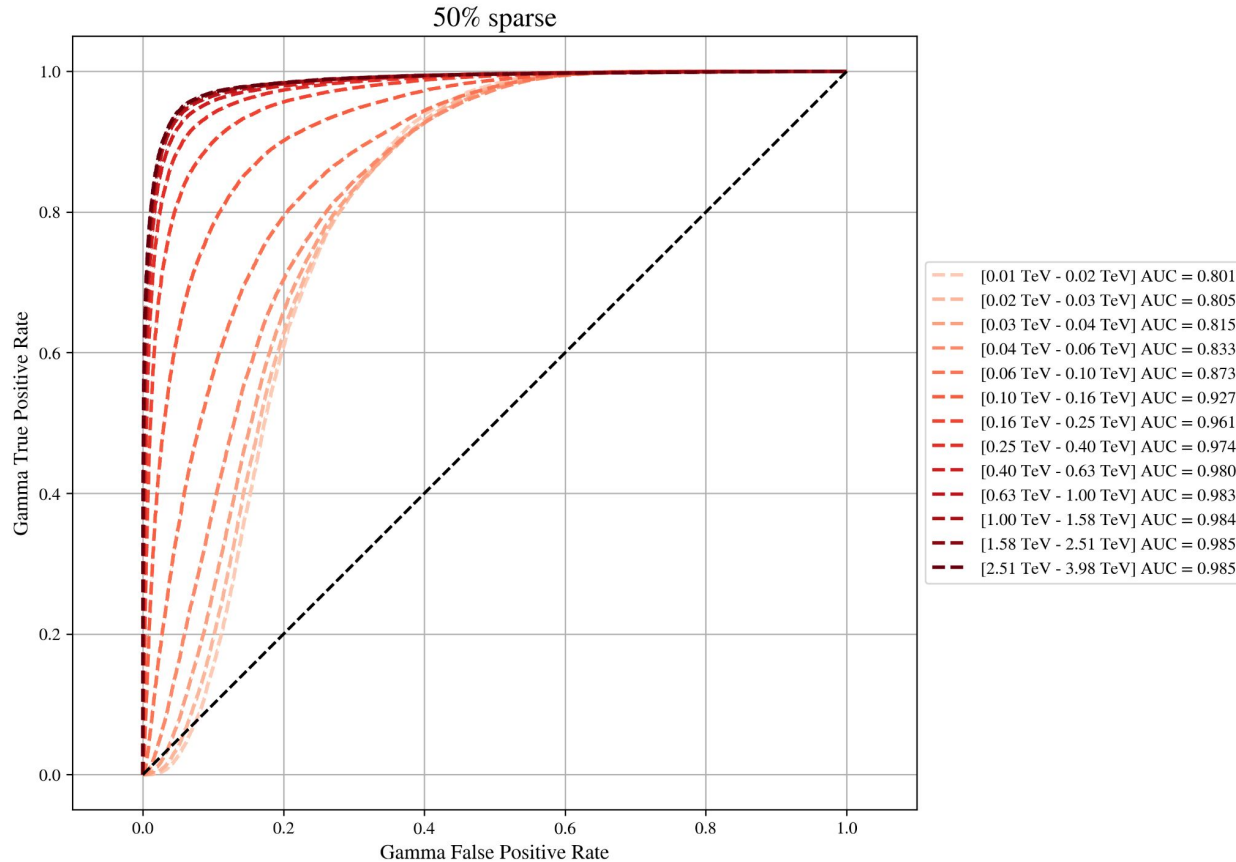
Resultados: Tarea de clasificación



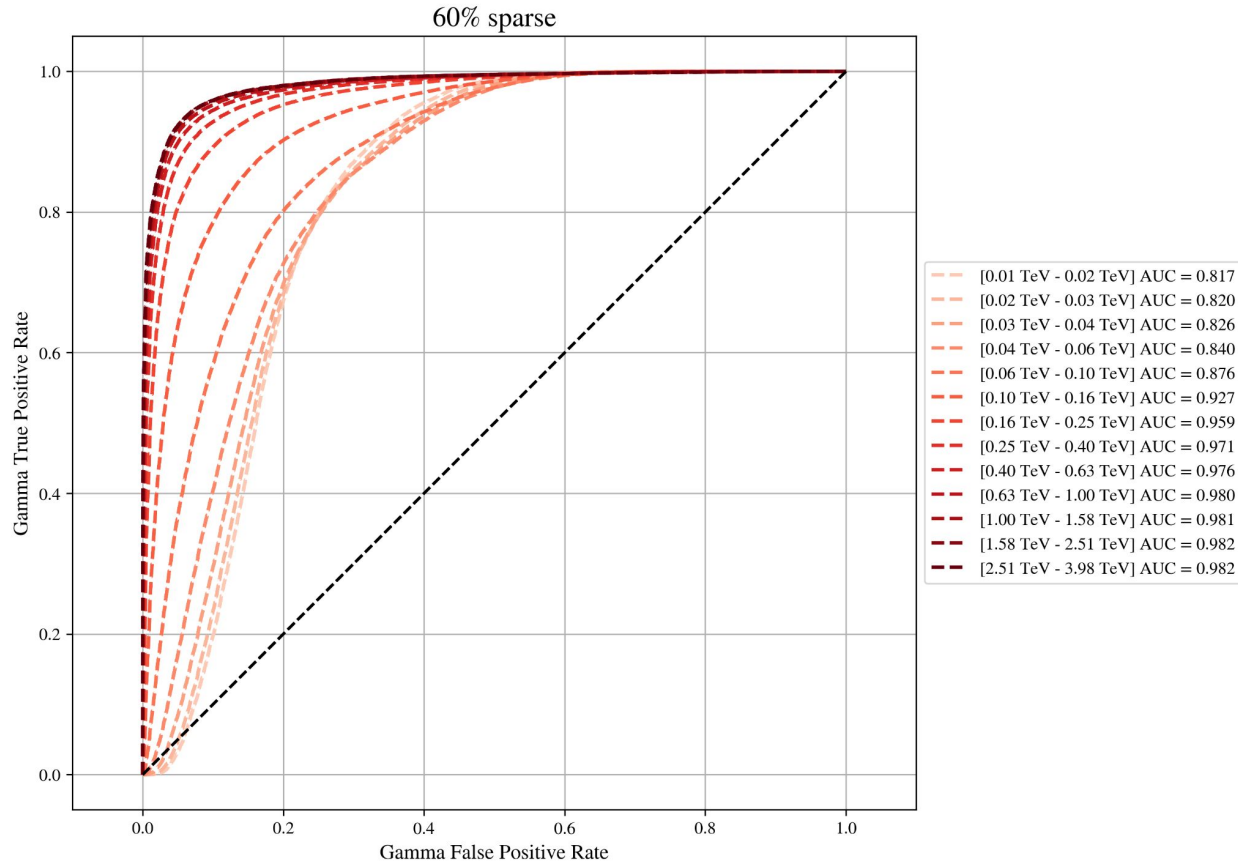
Resultados: Tarea de clasificación



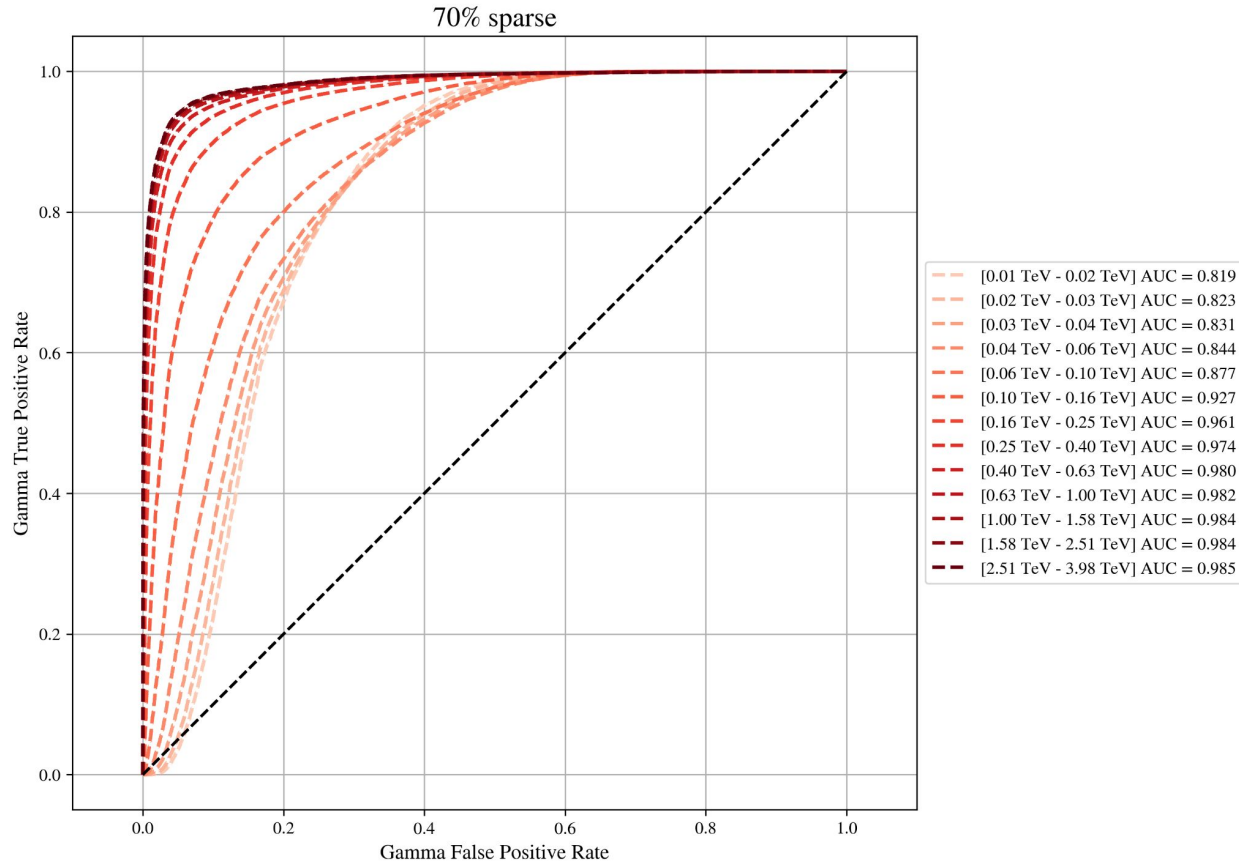
Resultados: Tarea de clasificación



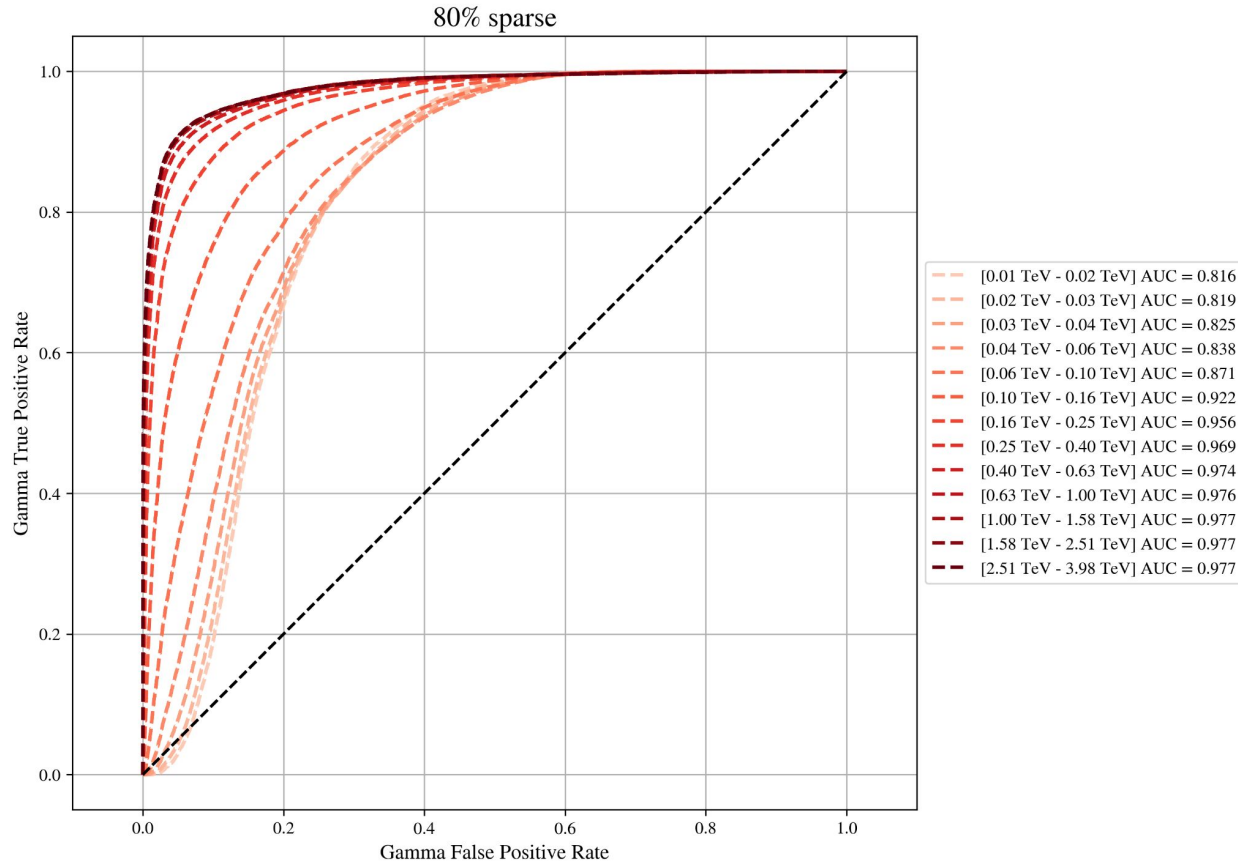
Resultados: Tarea de clasificación



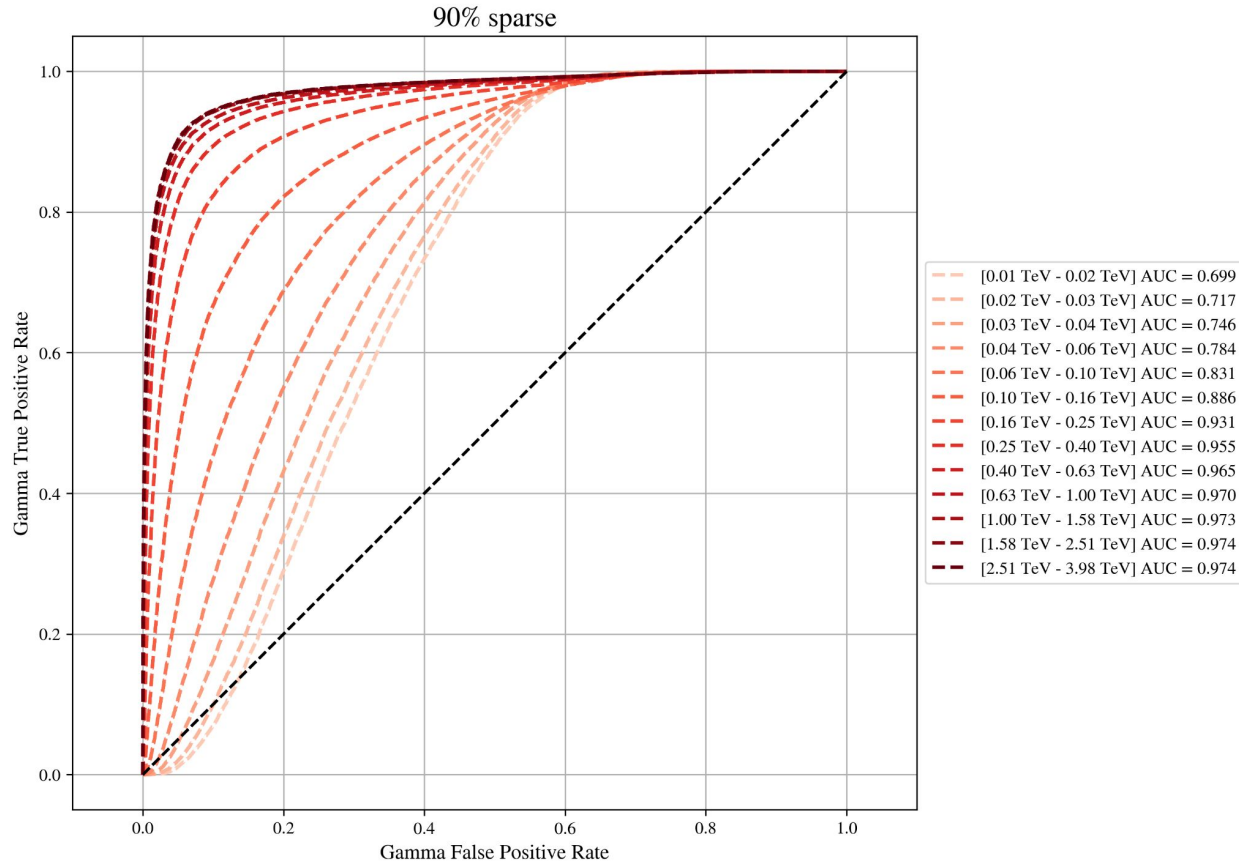
Resultados: Tarea de clasificación



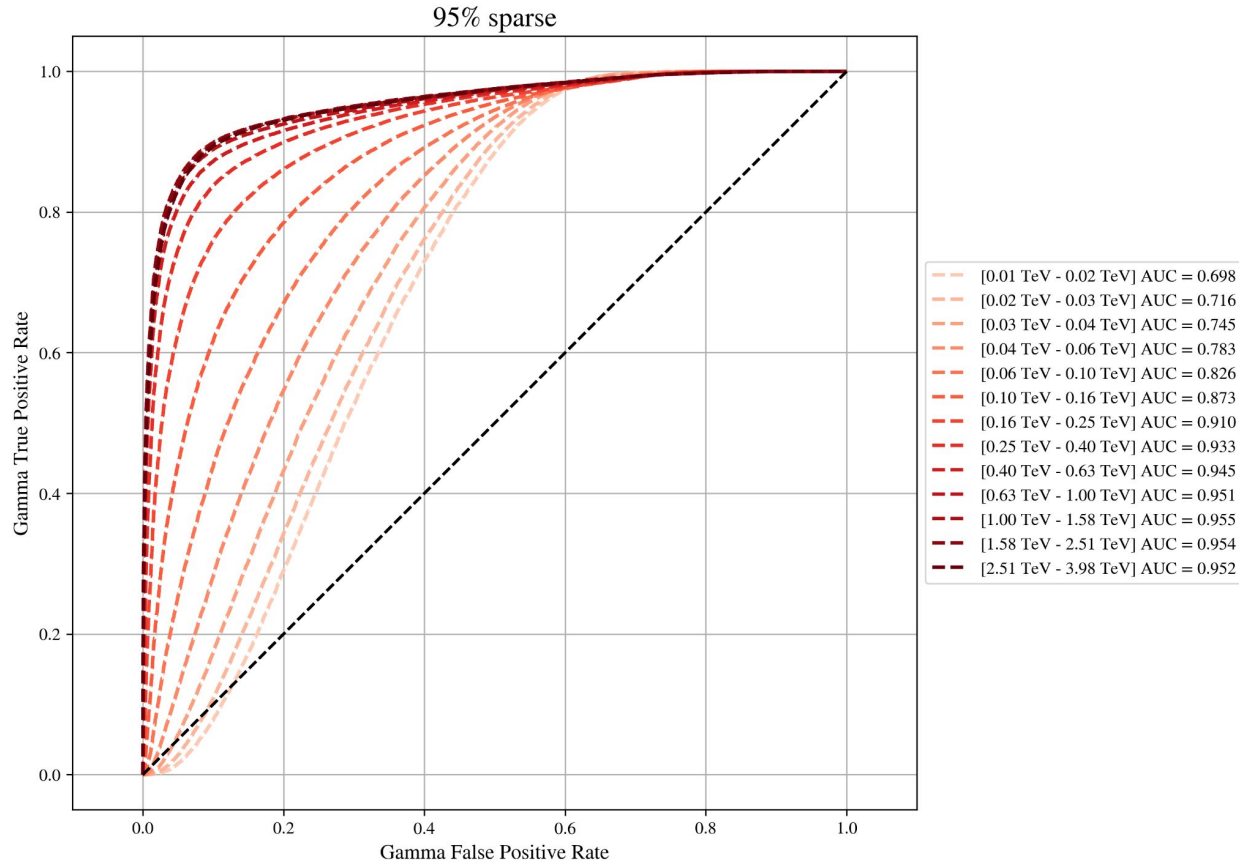
Resultados: Tarea de clasificación



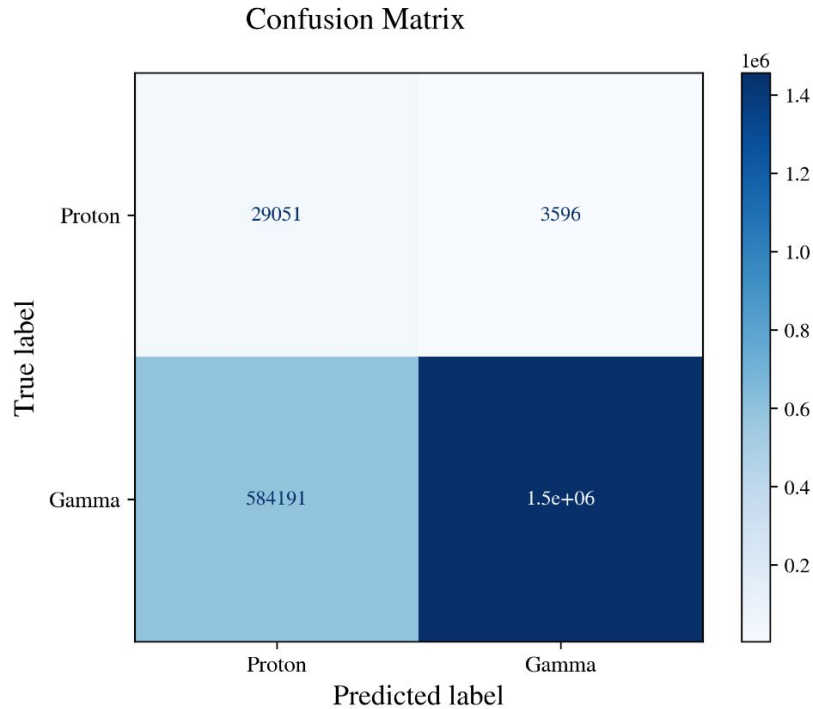
Resultados: Tarea de clasificación



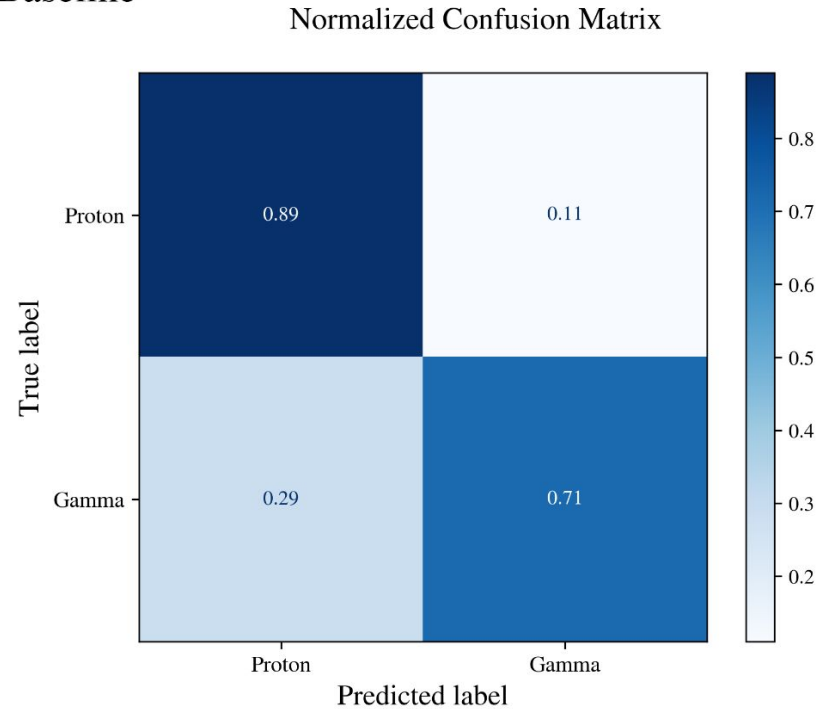
Resultados: Tarea de clasificación



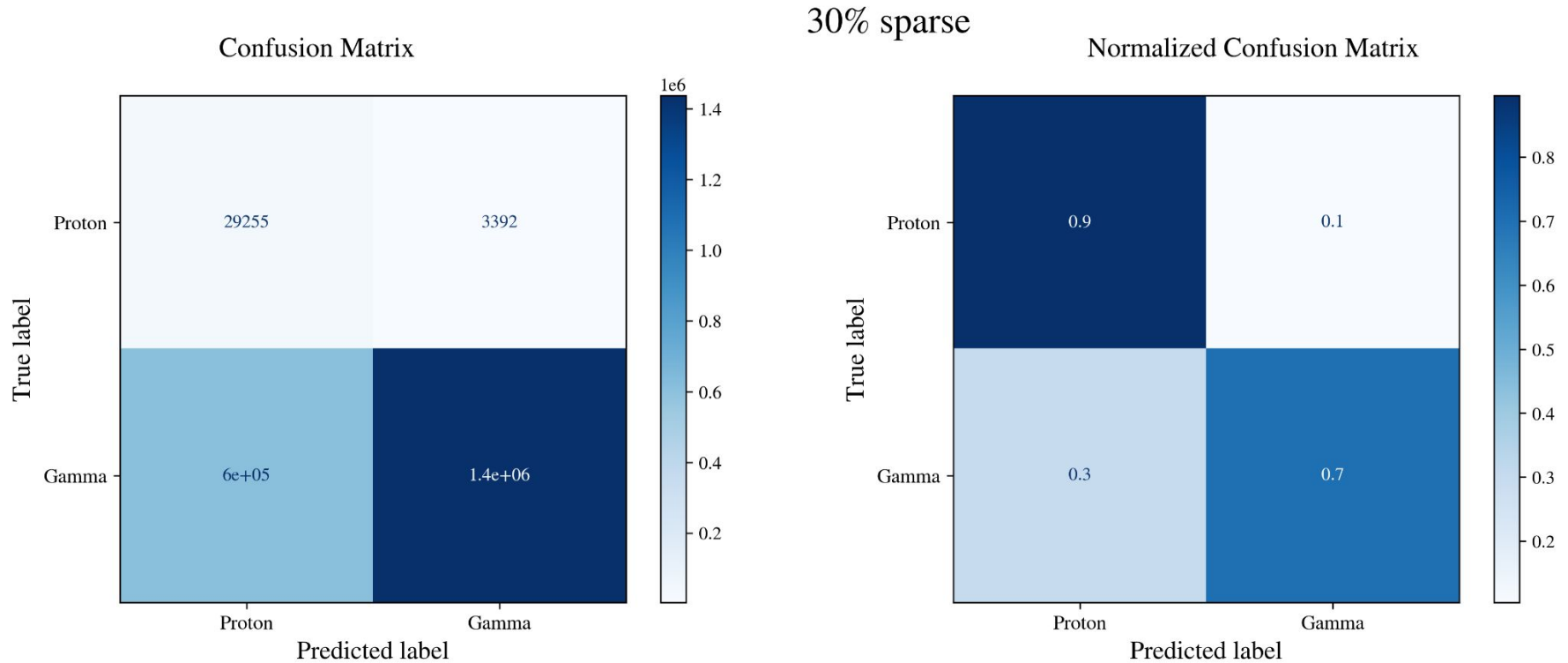
Resultados: Tarea de clasificación



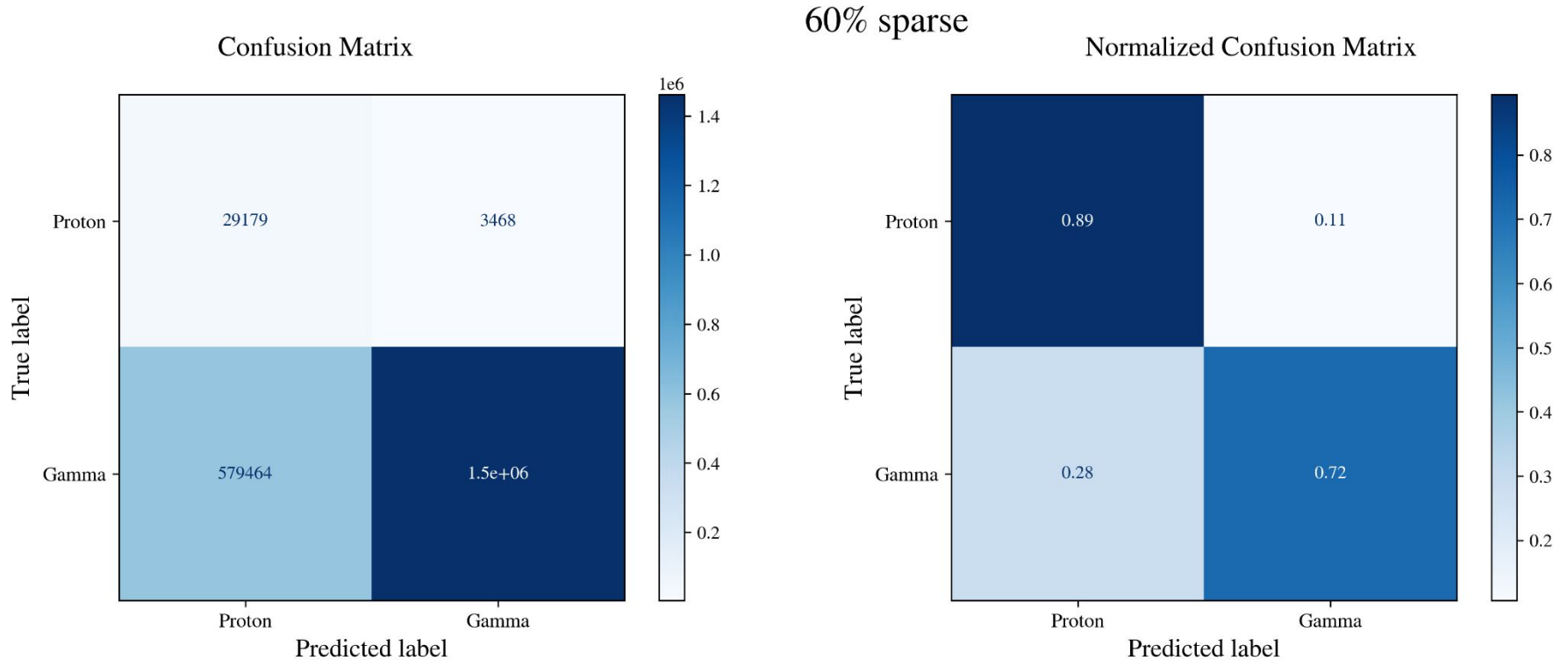
Baseline



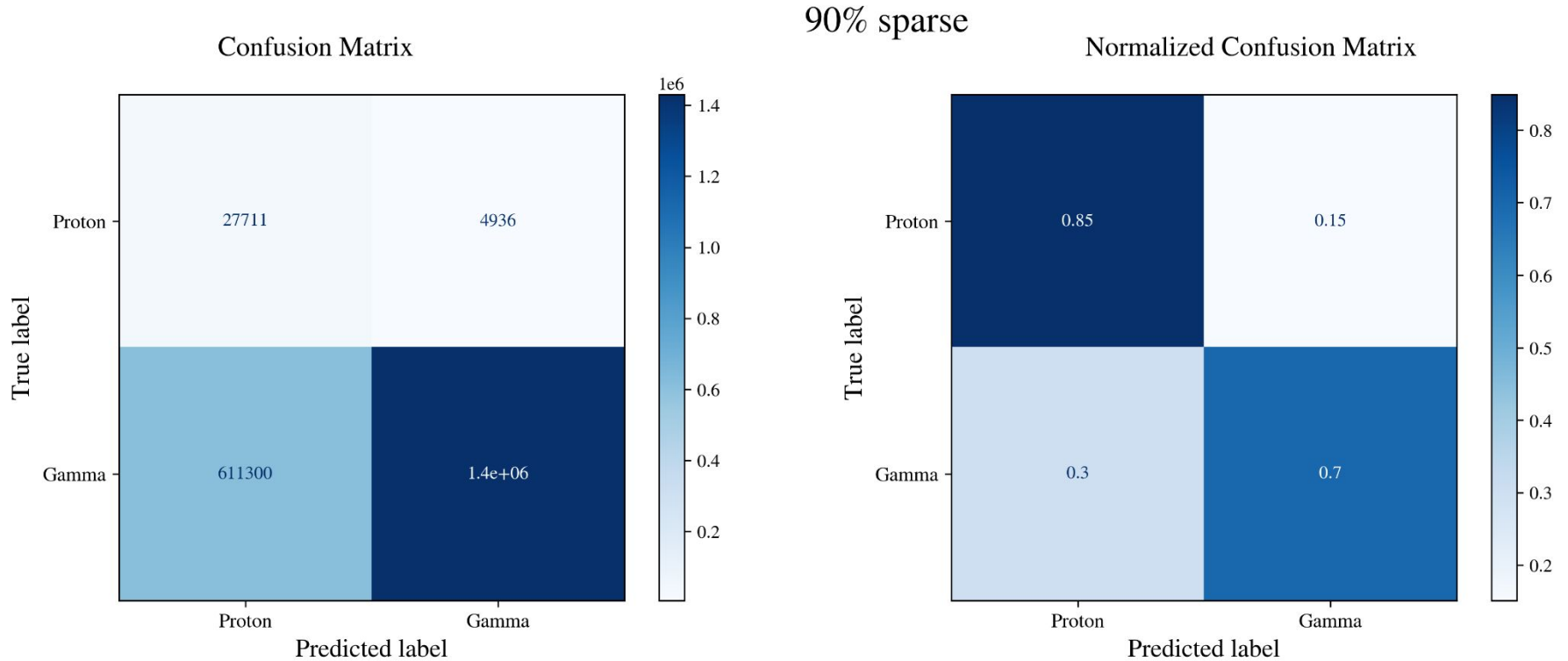
Resultados: Tarea de clasificación



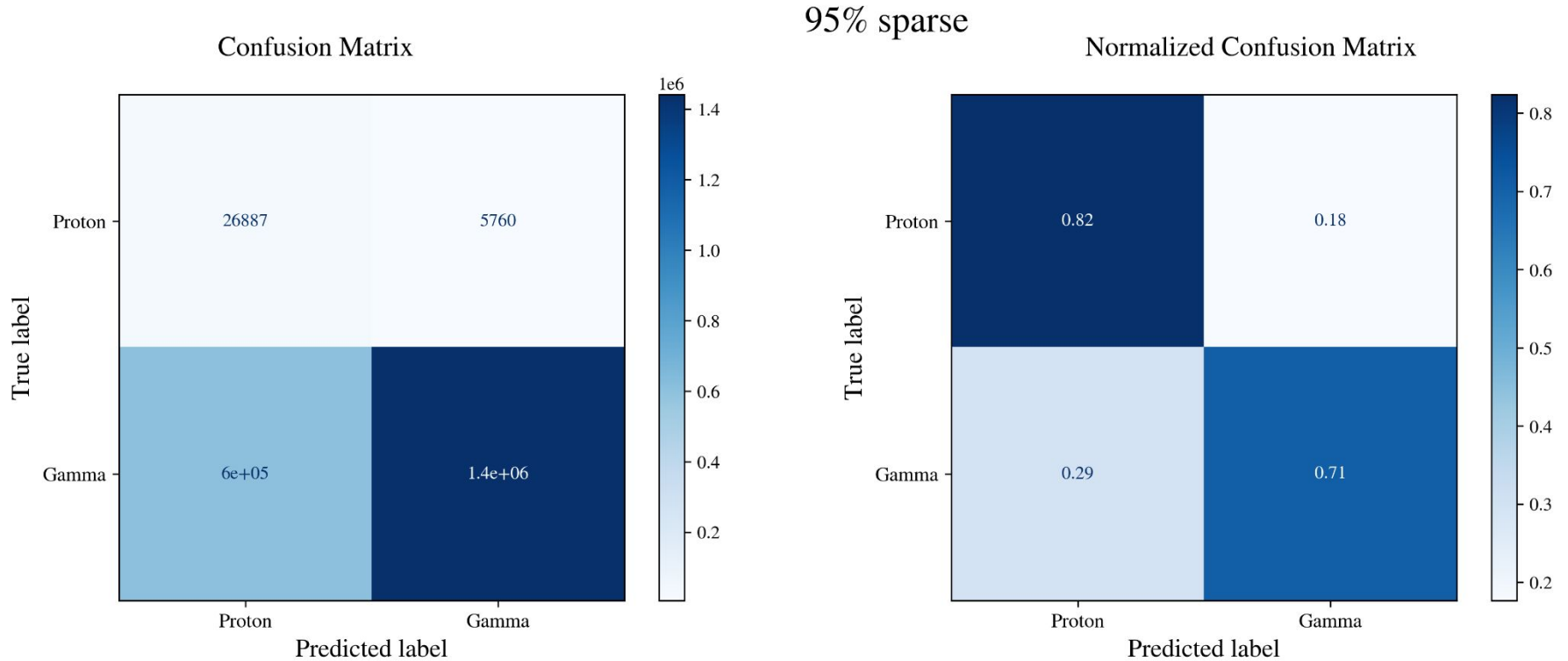
Resultados: Tarea de clasificación



Resultados: Tarea de clasificación



Resultados: Tarea de clasificación



Resultados: Tarea de clasificación

Datos:

Train:

- Gamma: 1,039,893
- Hadrónicos: 1,042,951

10% del conjunto de train se usa para validación.

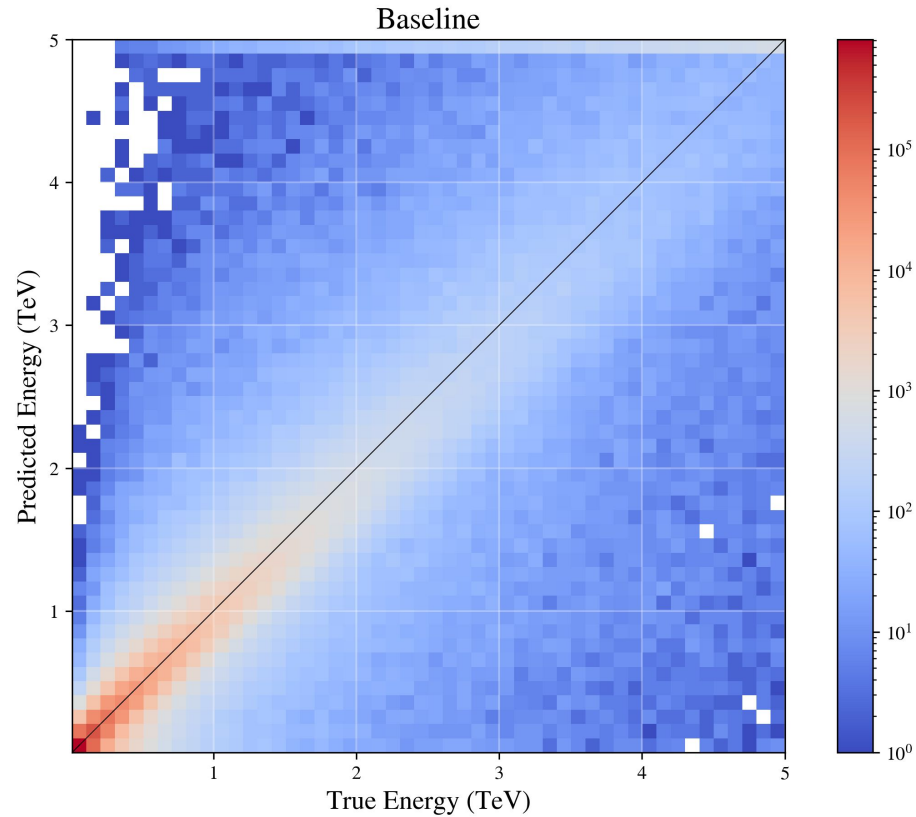
Test:

- Gamma: 2,041,124
- Hadrónicos: 32,647

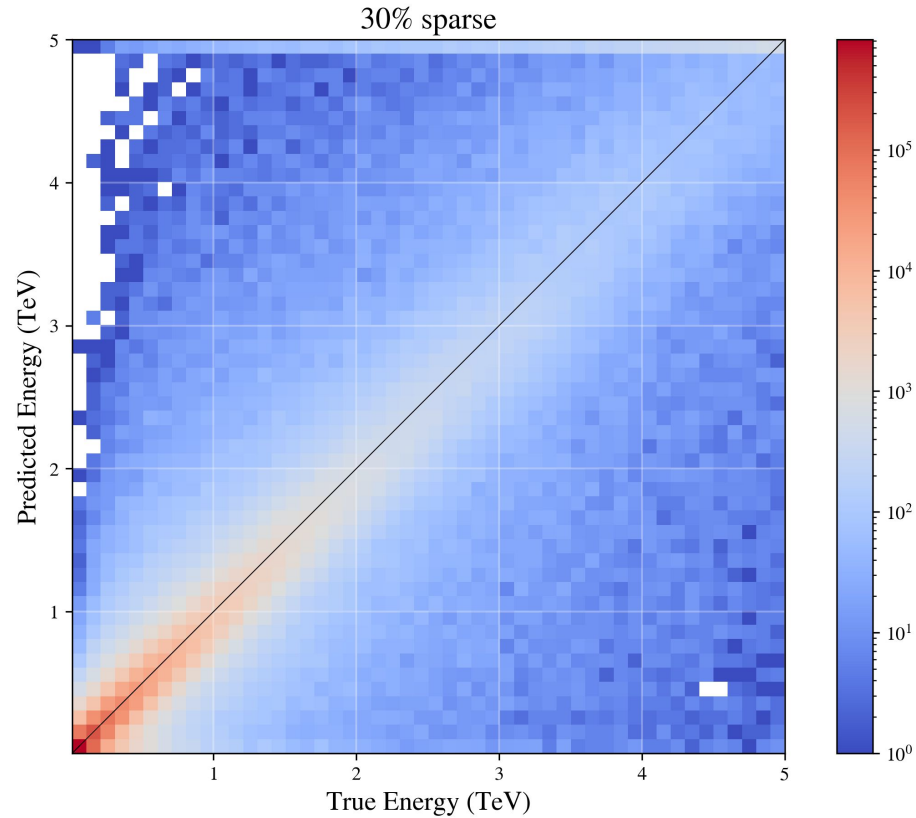
Resultados: Tarea de clasificación

	Metric	Baseline	30% sparse	40% sparse	50% sparse	60% sparse	70% sparse	80% sparse	90% sparse	95% sparse
0	Accuracy	71.66%	70.72%	72.20%	70.56%	71.89%	71.25%	70.71%	70.28%	70.78%
1	Precision (Class 0)	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04
2	Precision (Class 1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	Recall (Class 0)	0.89	0.90	0.90	0.90	0.89	0.91	0.89	0.85	0.82
4	Recall (Class 1)	0.71	0.70	0.72	0.70	0.72	0.71	0.70	0.70	0.71
5	F1 Score	0.83	0.83	0.84	0.82	0.83	0.83	0.83	0.82	0.83
6	Support (Class 0)	32647.0	32647.0	32647.0	32647.0	32647.0	32647.0	32647.0	32647.0	32647.0
7	Support (Class 1)	2041124.0	2041124.0	2041124.0	2041124.0	2041124.0	2041124.0	2041124.0	2041124.0	2041124.0
8	AUC	0.91	0.91	0.92	0.91	0.91	0.92	0.91	0.88	0.86

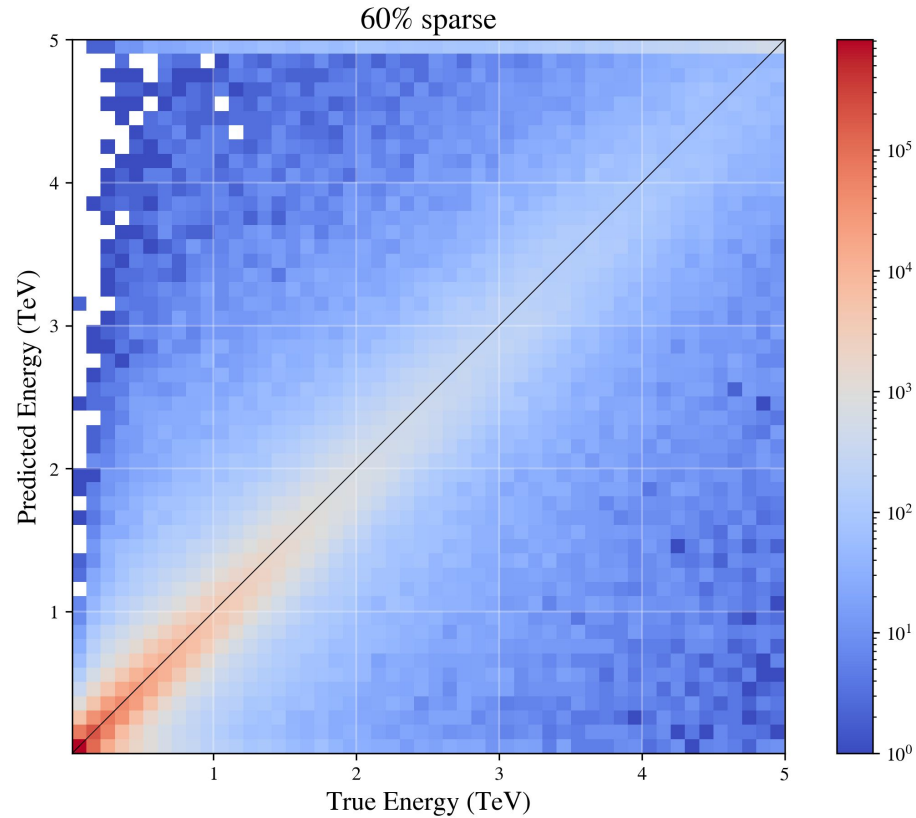
Resultados: Tarea de regresión



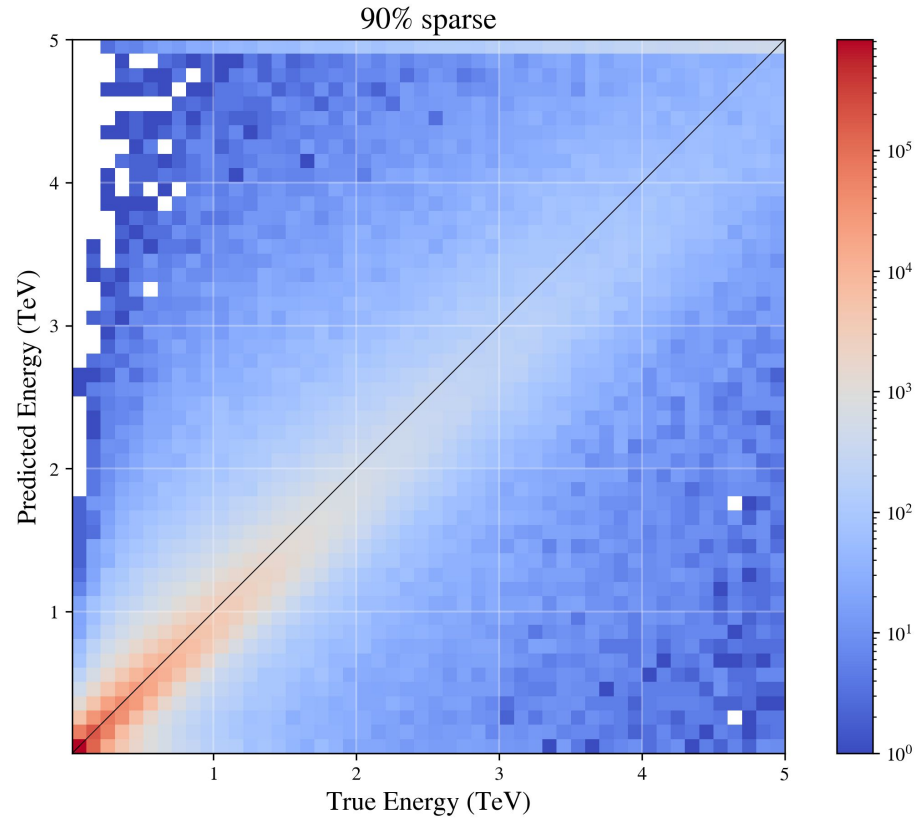
Resultados: Tarea de regresión



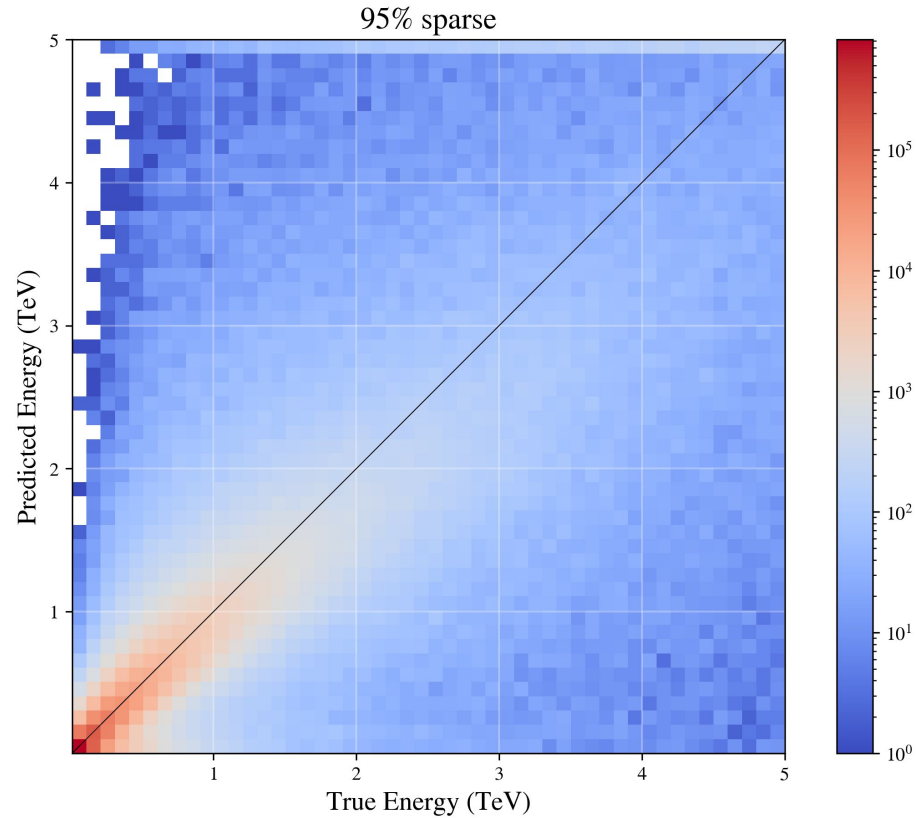
Resultados: Tarea de regresión



Resultados: Tarea de regresión



Resultados: Tarea de regresión

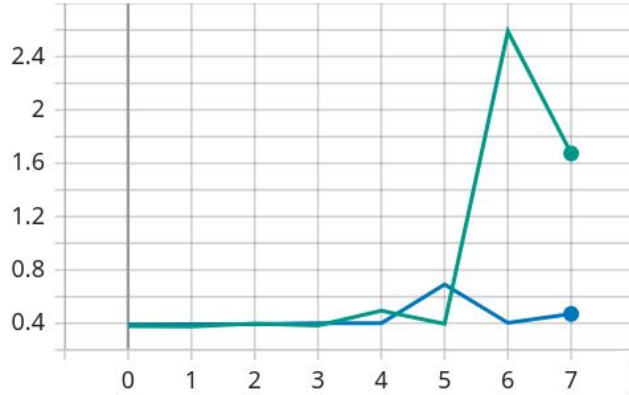


Resultados: Tarea de regresión

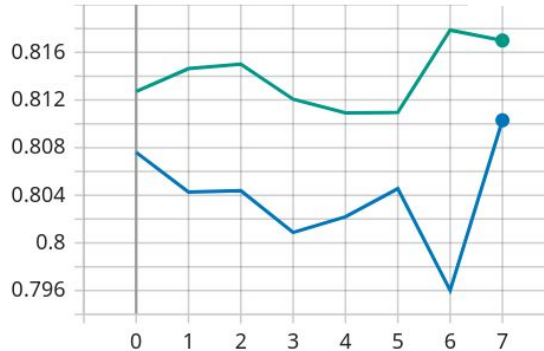
Model size in (MB)	Baseline	30%	40%	50%	60%	70%	80%	90%	95%
Type of shower	67.56	18.77	16.82	14.67	12.59	10.19	7.80	5.15	3.68
Direction	69.05	18.74	16.78	14.72	12.53	10.23	7.77	5.14	3.62
Energy	66.10	18.70	16.74	14.66	12.49	10.19	7.81	5.13	3.68

Monitorización clasificación: TensorBoard

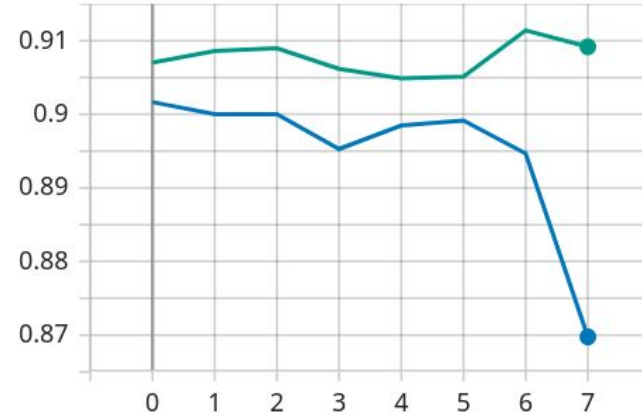
epoch_loss
tag: epoch_loss



epoch_accuracy
tag: epoch_accuracy

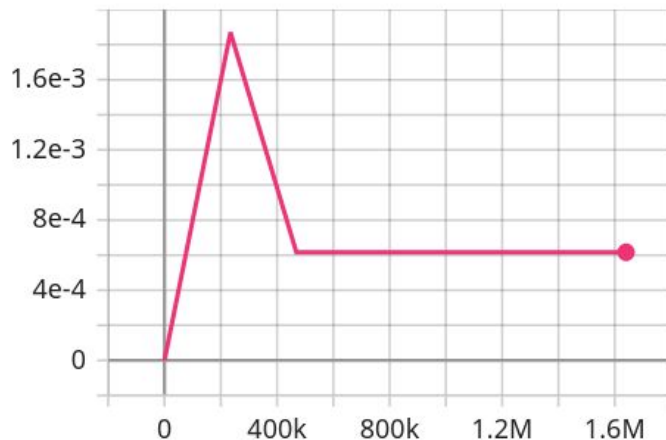


epoch_auc
tag: epoch_auc

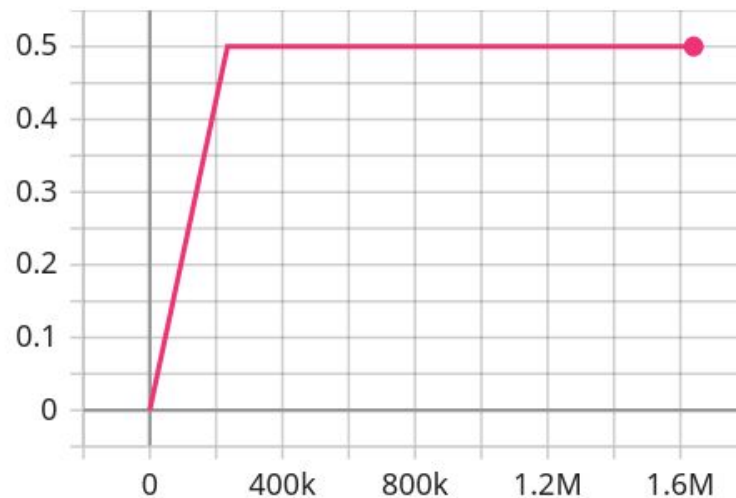


Monitorización clasificación: TensorBoard

prune_low_magnitude_type/threshold:0/
threshold
tag: prune_low_magnitude_type/threshold:0/
threshold

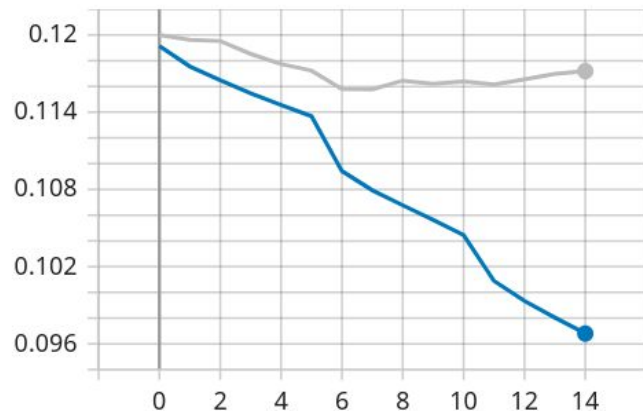


prune_low_magnitude_type/mask:0/sparsity
tag: prune_low_magnitude_type/mask:0/sparsity

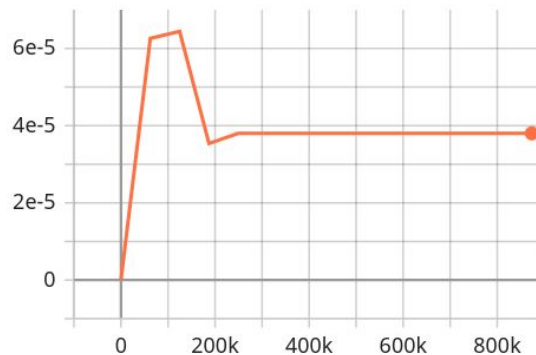


Monitorización: TensorBoard

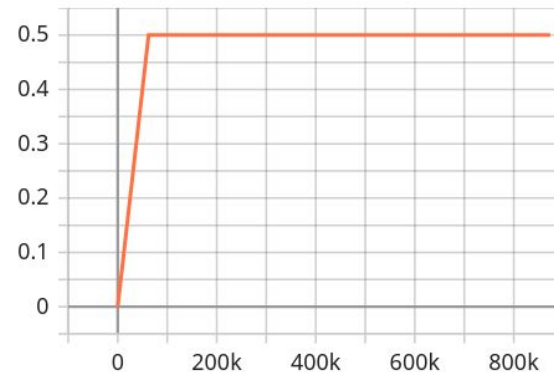
epoch_loss
tag: epoch_loss



prune_low_magnitude_cameradirection/
threshold:0/threshold
tag: prune_low_magnitude_cameradirection/
threshold:0/threshold

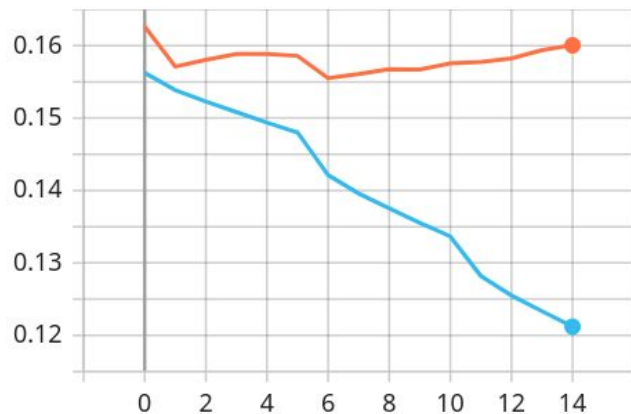


prune_low_magnitude_cameradirection/mask:0/
sparsity
tag: prune_low_magnitude_cameradirection/
mask:0/sparsity

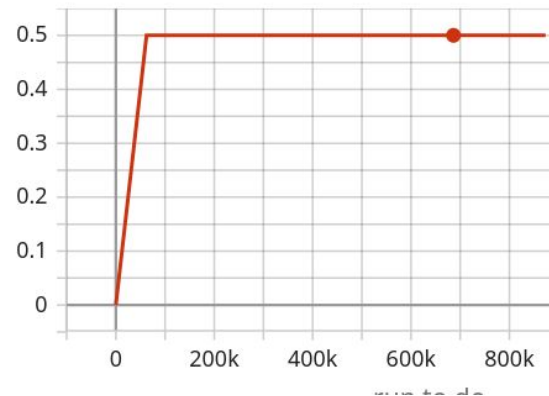


Monitorización: TensorBoard

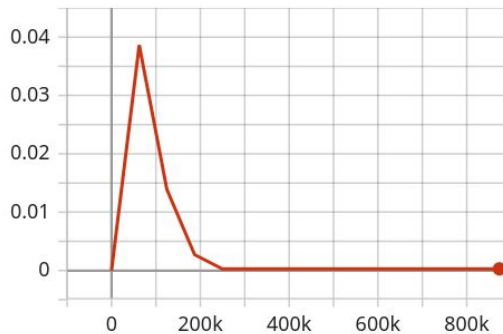
epoch_loss
tag: epoch_loss



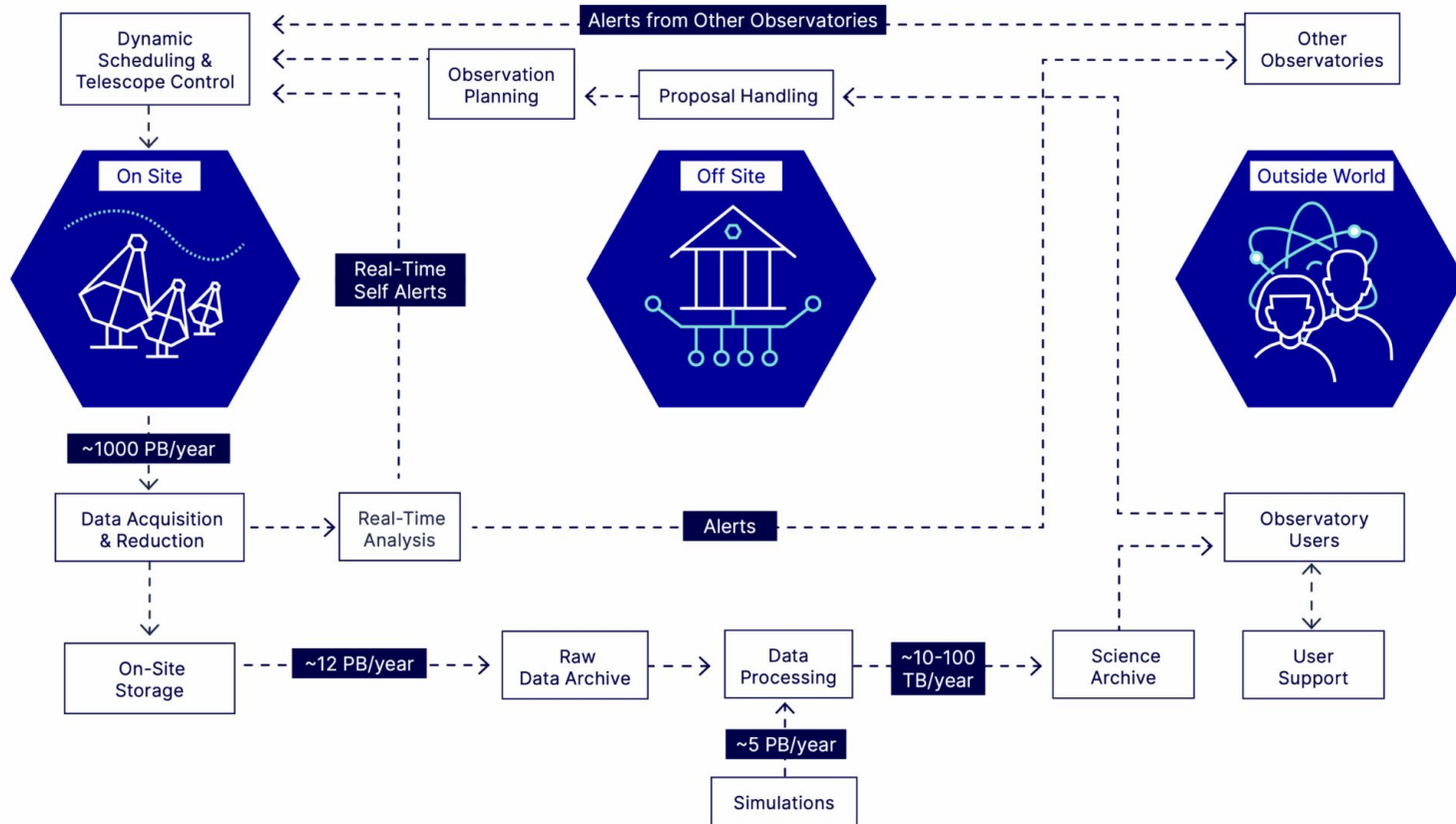
prune_low_magnitude_energy/mask:0/sparsity
tag: prune_low_magnitude_energy/mask:0/
sparsity



prune_low_magnitude_energy/threshold:0/
threshold
tag: prune_low_magnitude_energy/threshold:0/
threshold



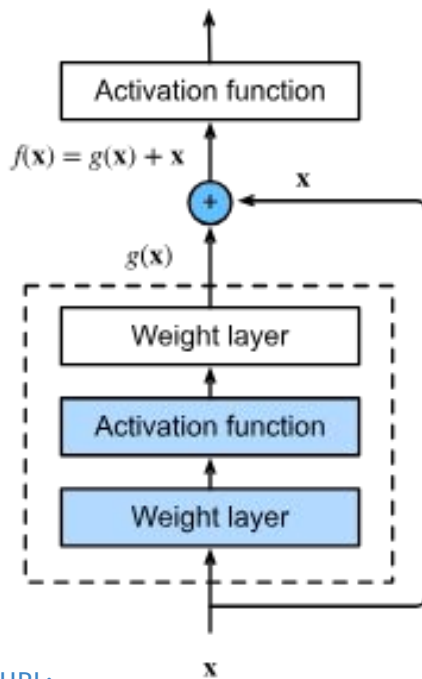
Data Rates



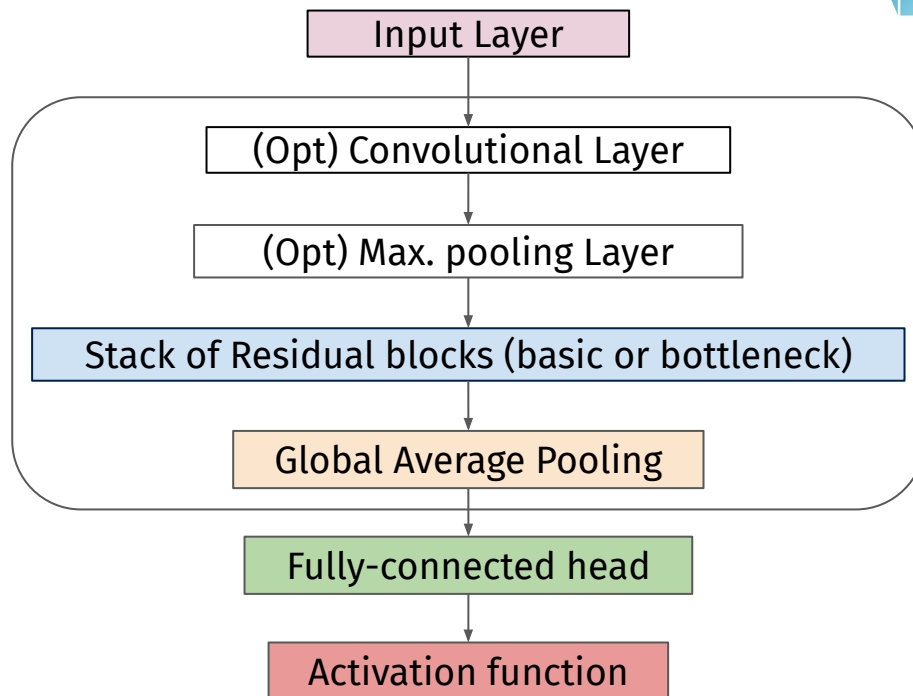
CTLearn: Arquitectura del Modelo backup



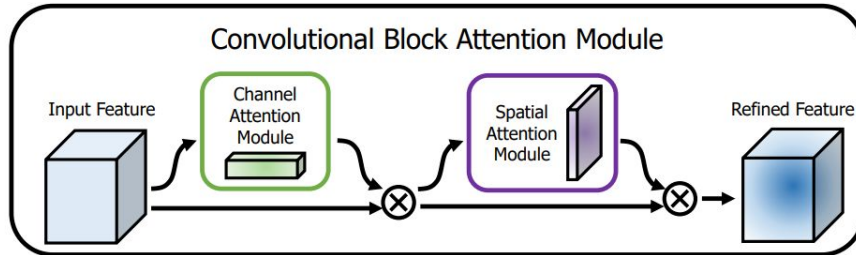
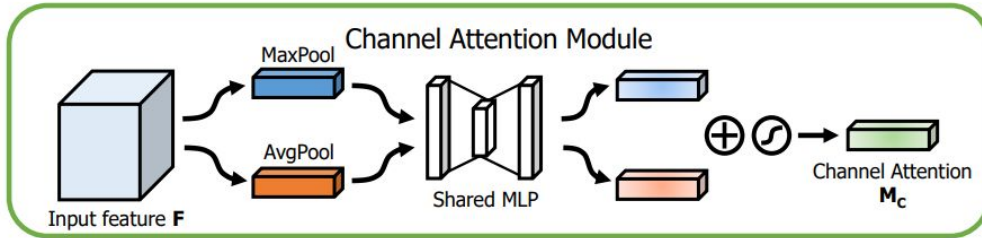
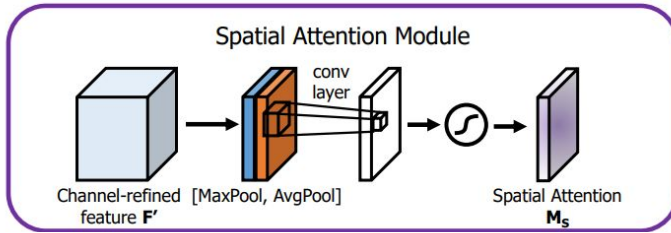
Bloque Residual

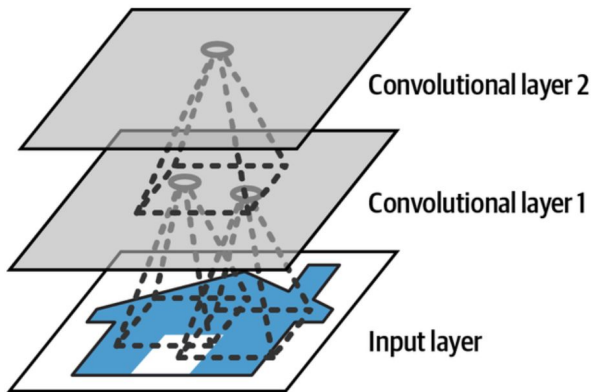


ResNet



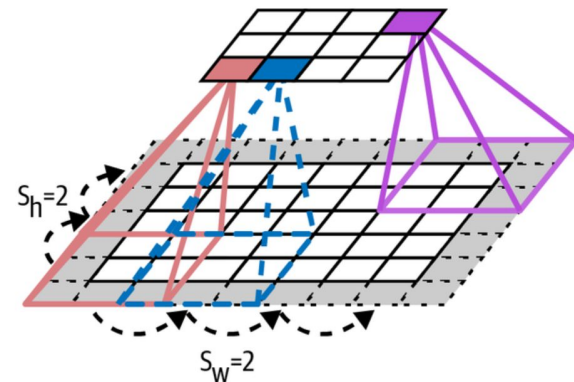
Mecanismo de atención squeeze-excite dual





Geron (2022)

Red convolutional



Geron (2022)

