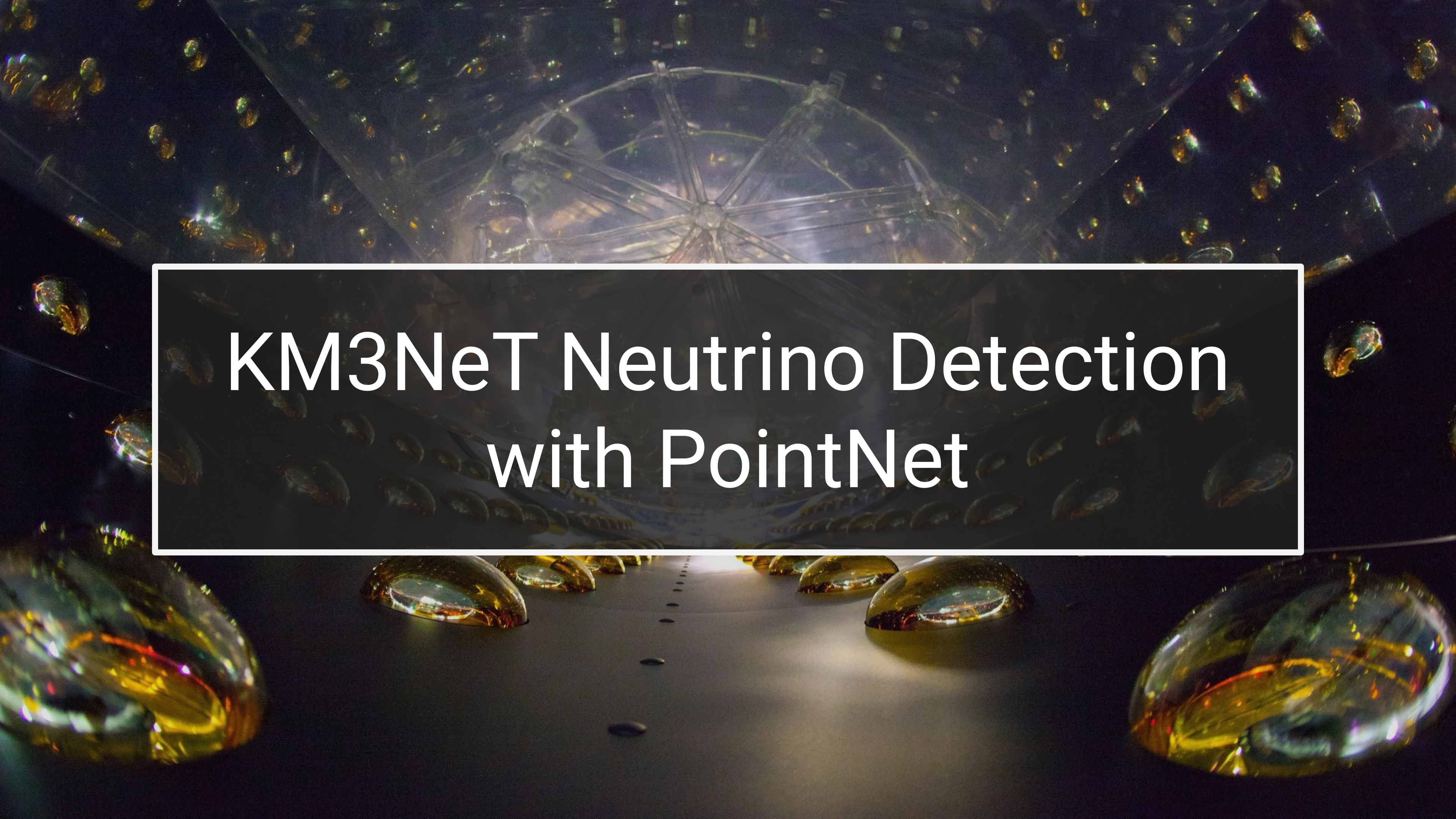
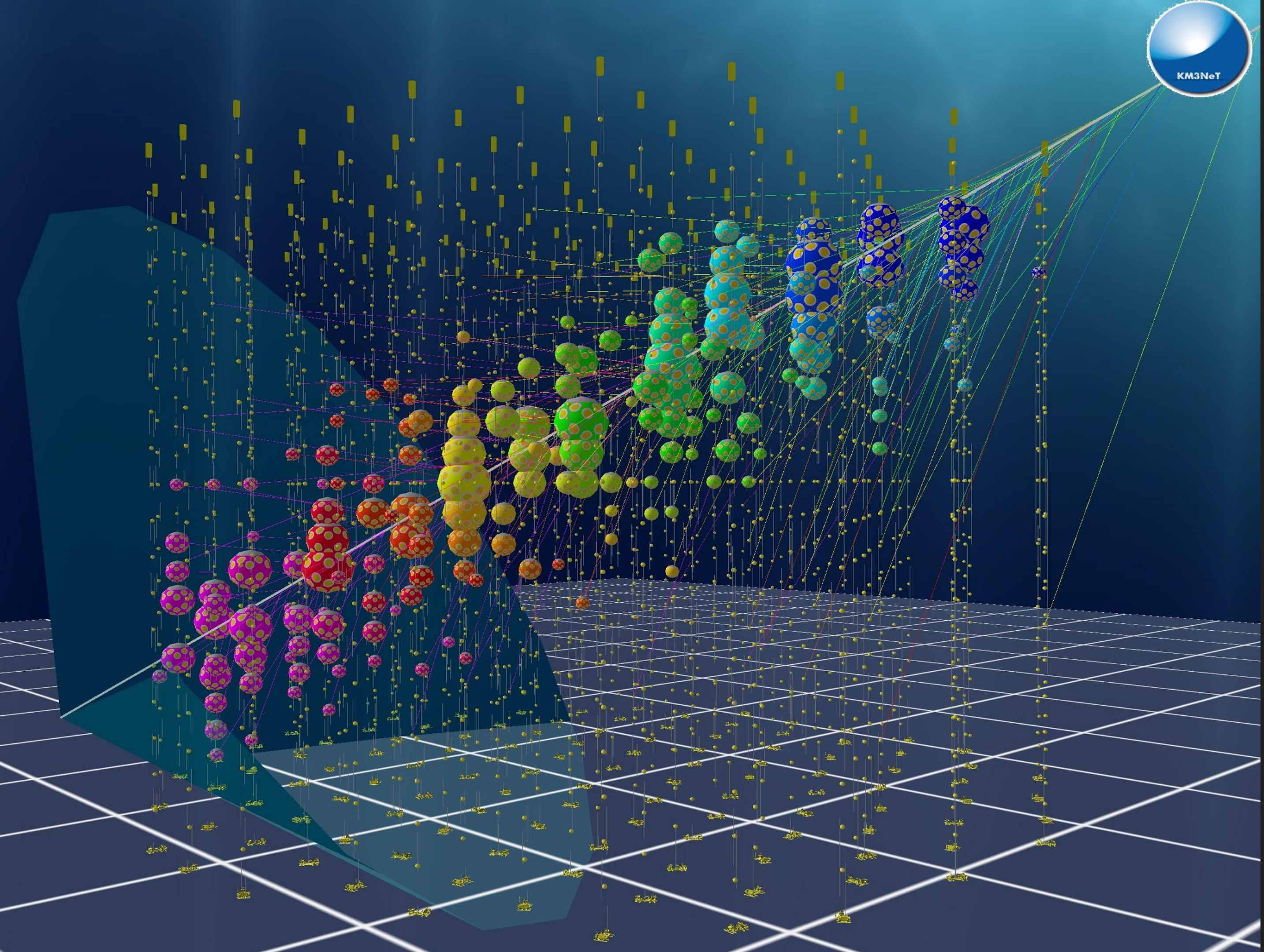


KM3NeT Neutrino Detection with PointNet

Background

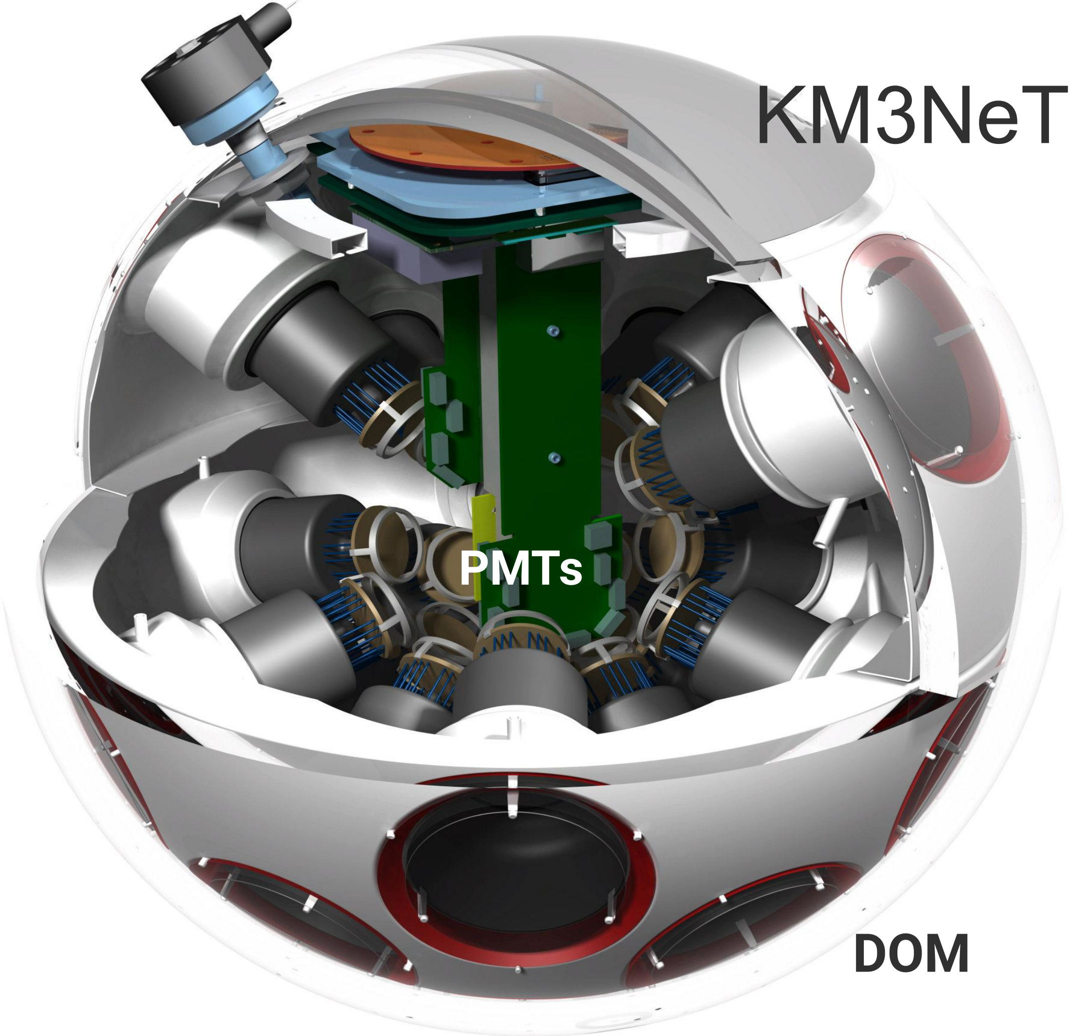


Neutrinos

⬇

Cherenkov
Light

KM3NeT



- Event Triggers - L0, L1, L2 [1]
- GPU Pipeline [2]
- CNN Experiment [1,3]

Motivation & Research Contributions

Motivation

- Assess state-of-the-art architectures
- Improve data acquisition for KM3NeT
- Shortcomings of CNNs [4, 5, 6]

Contributions

- Novel 3D mesh representation
- First known use of PointNet
- Improvement over L1 Trigger

Research Questions

RQ 1.0 Can PointNet classify noise timeslices and event timeslices?

RQ 1.1 Can PointNet obtain a Recall score of 0.9 for the positive class?

RQ 2.0 Can the KM3NeT dataset be effectively represented using 3D meshes?

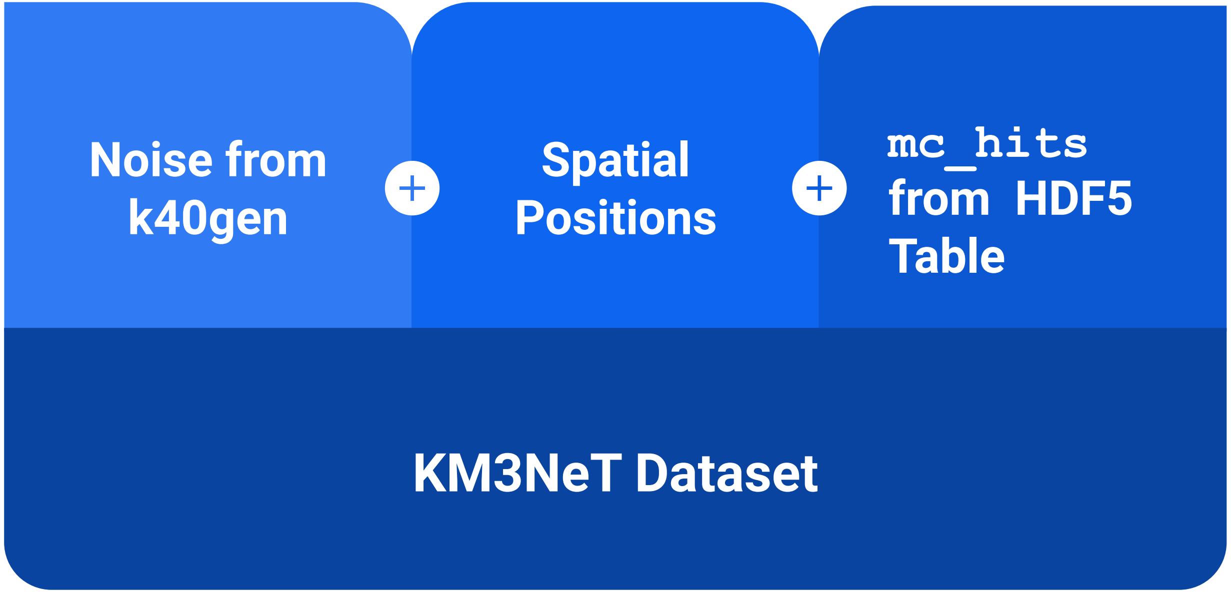
RQ 2.1 Which meshing algorithm would be suitable for representing data?

RQ 3.0

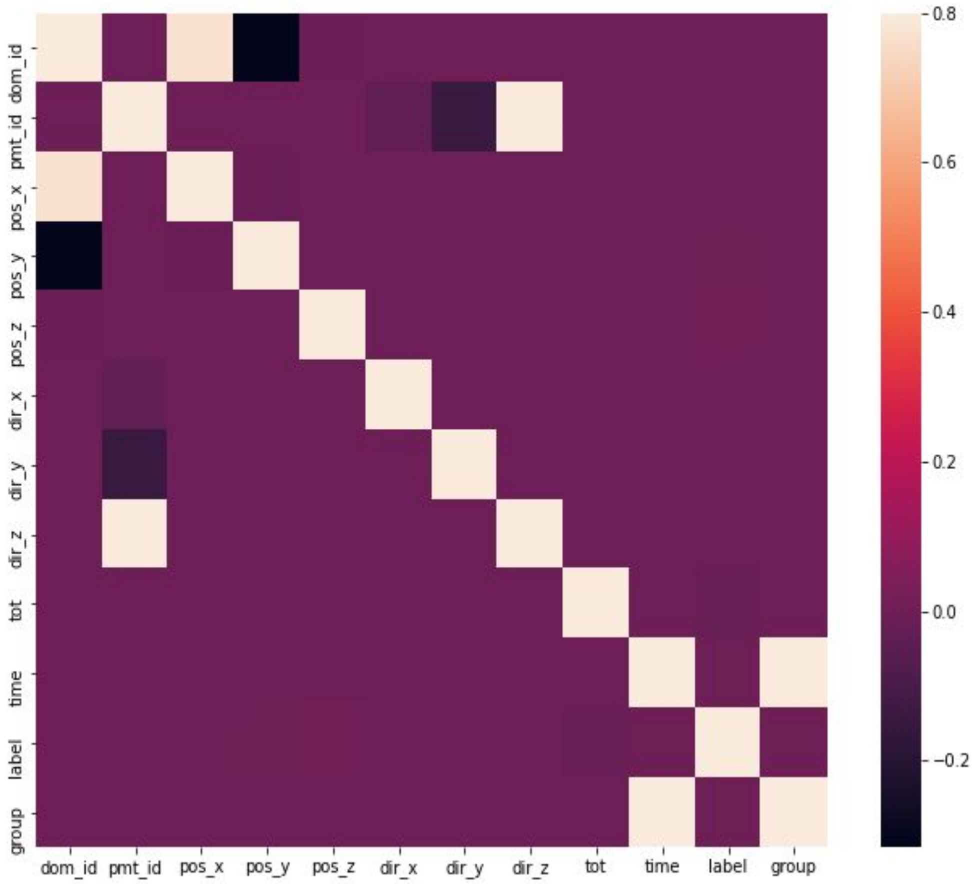
Can PointNet be extended for energy inference from events?

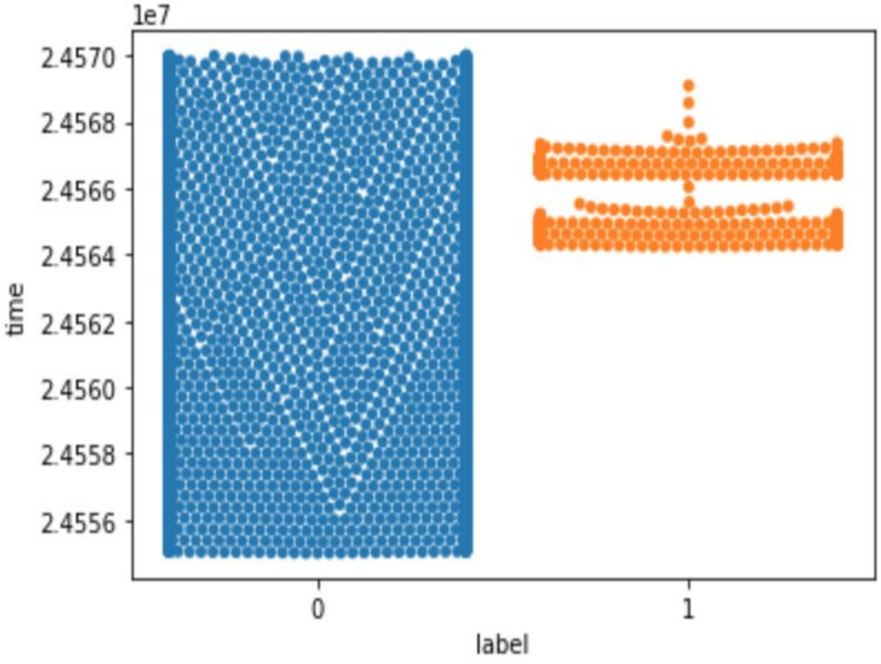
KM3NeT Dataset

Data Preparation

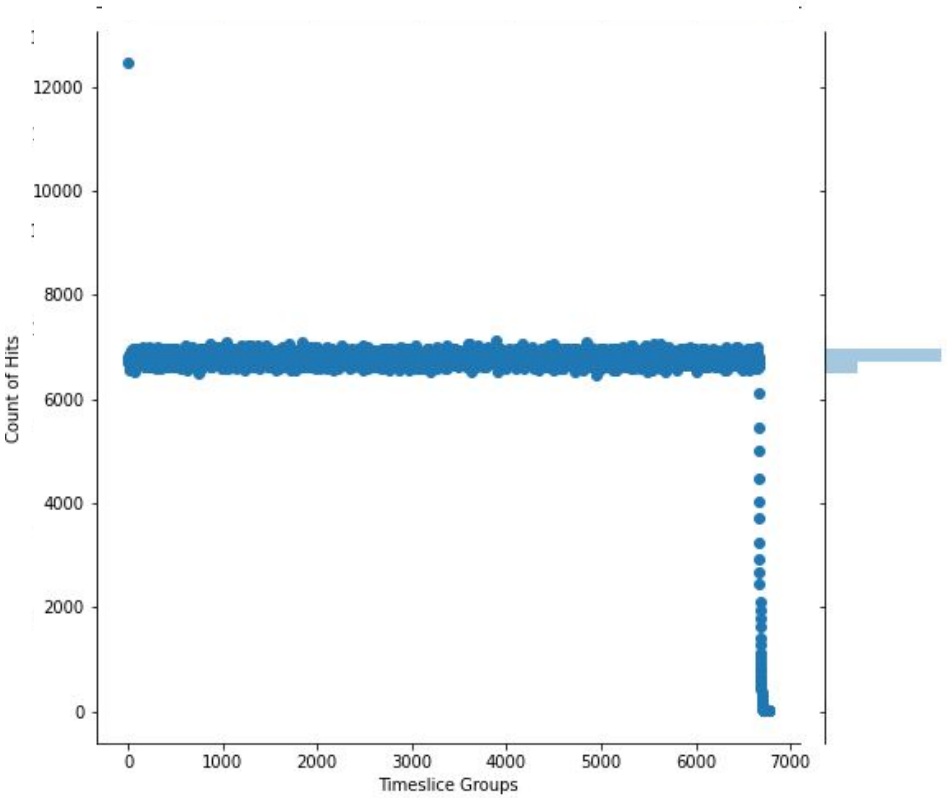


Attribute	Description
`dom_id`	[Unique ID for sensor module.]
`pmt_id`	[Unique ID for photomultiplier (PMT) tubes within DOMs.]
`pos_x, pos_y, pos_z`	[Spatial coordinates (in meters) of hit within the detector.]
`dir_x, dir_y, dir_z`	[Direction of PMT tubes within DOMs to look for Cherenkov Light.]
`tot`	[Time-over-threshold (ToT) indicates the amount of light transformed to charge which is interpreted as the length of the square wave pulse over a given threshold (26).]
`time`	[Time at which the hit was recorded.]
`label`	[0 or 1 class label indicating whether hit is from noise or event respectively.]
`group`	[Timeslice numbers starting from 0 for the purpose of identification.]

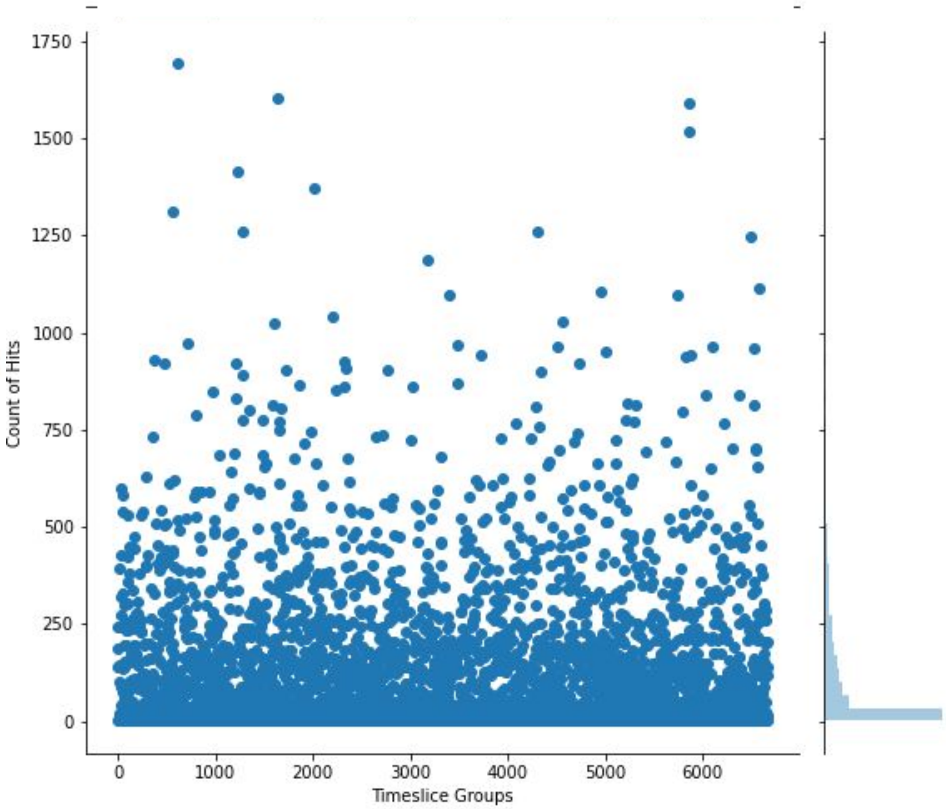

dom_id
pmt_id
pos_x
pos_y
pos_z
dir_x
dir_y
dir_z
tot
time
label
group
0.8
0.6
0.4
0.2
0.0
−0.2


1e7
2.4570
2.4568
2.4566
2.4564
2.4562
2.4560
2.4558
2.4556
time
0
1
label

Noise Timeslices

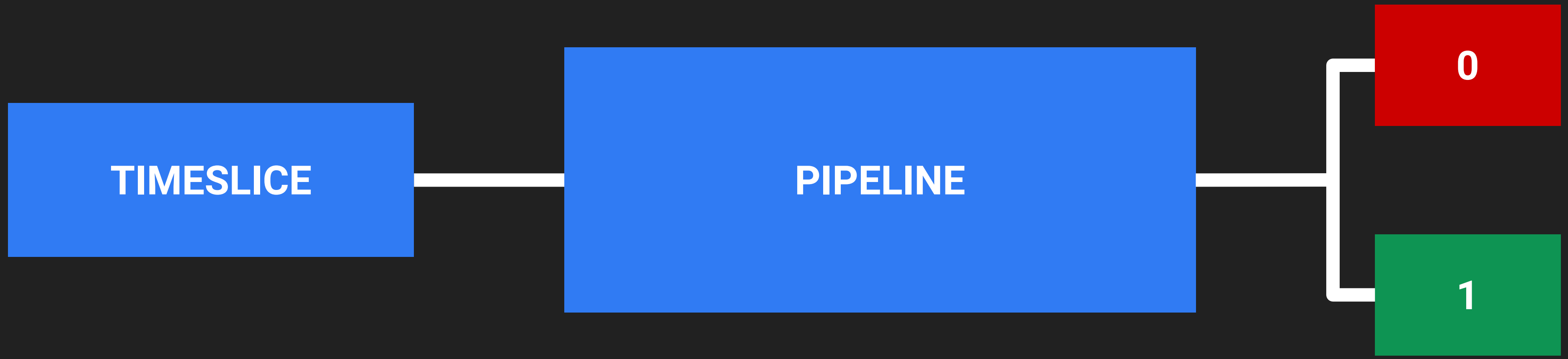


Event Timeslices



The Pipeline

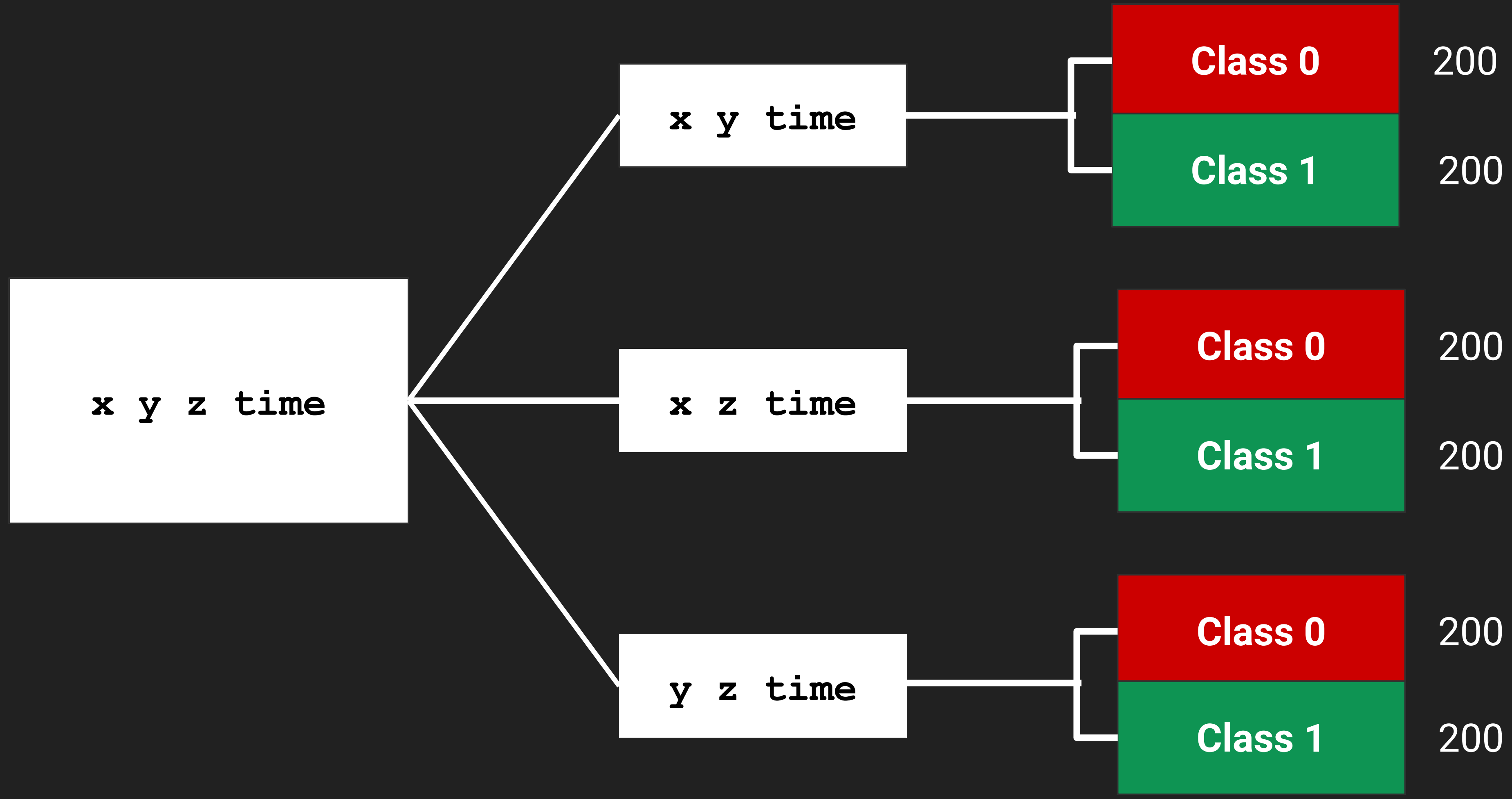
Design Goal



Noise Timeslices : `class_0`

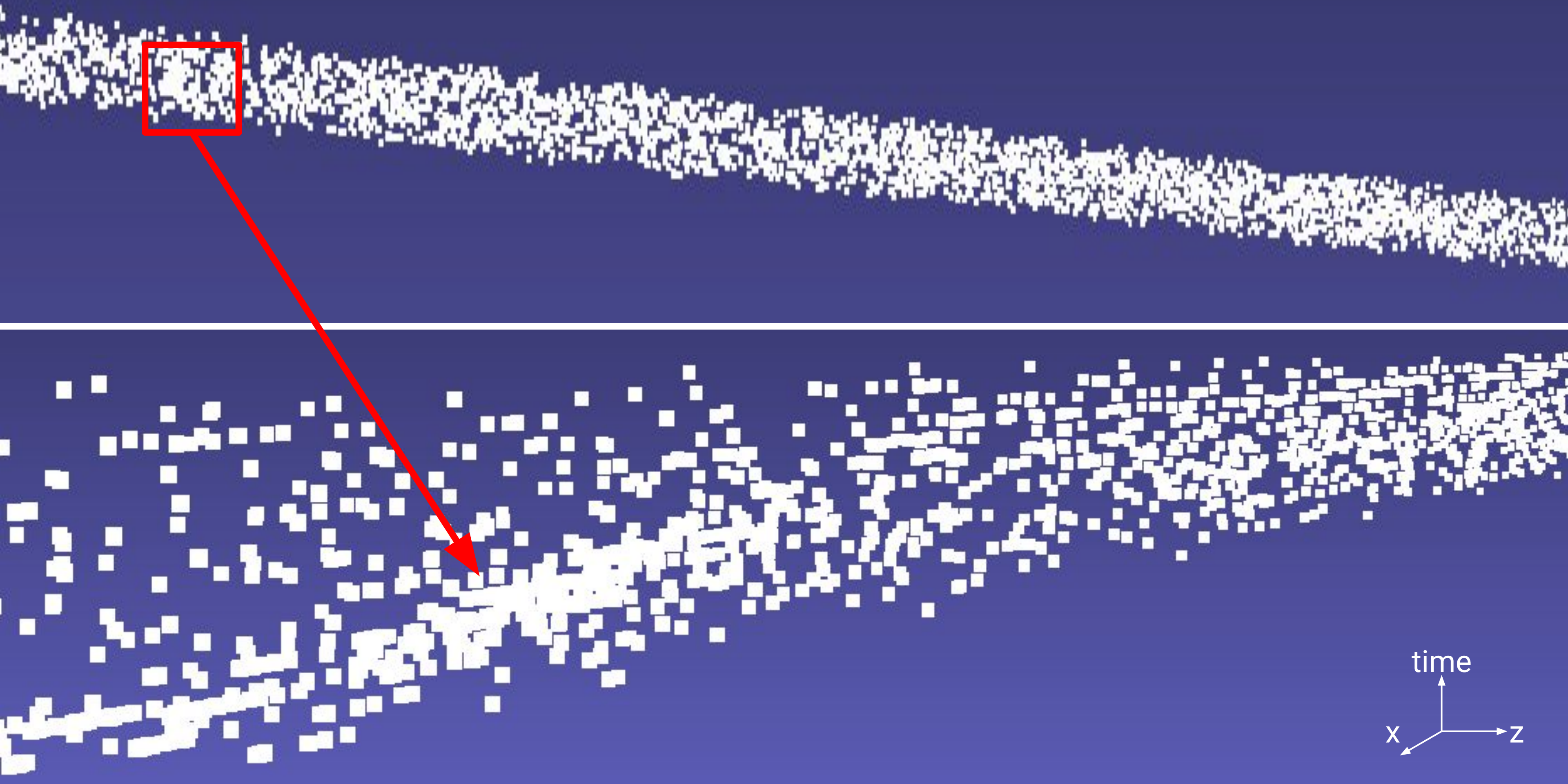
Event Timeslices : `class_1`

Step 1: 3D Coordinates


x y z time
x y time
x z time
y z time
Class 0
Class 1
200
200
Class 0
Class 1
200
200
Class 0
Class 1
200
200

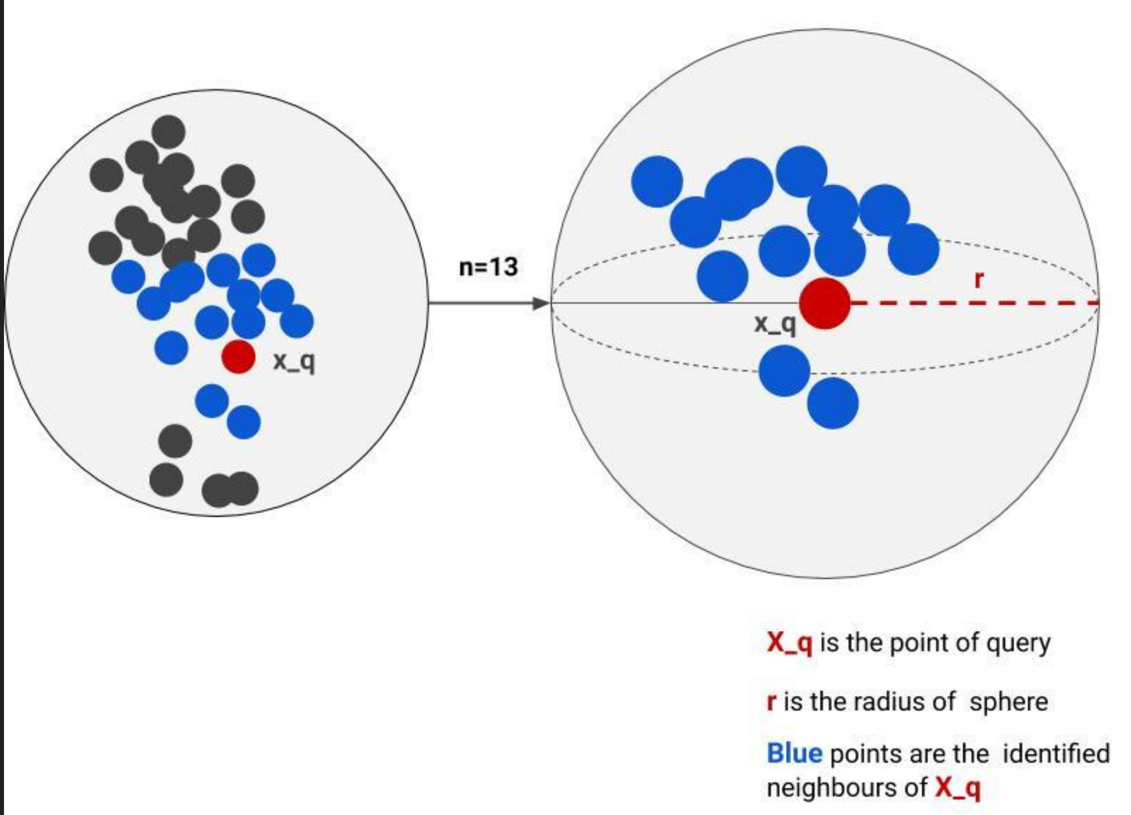
Step 2: Feature Engineering

Identified Problem

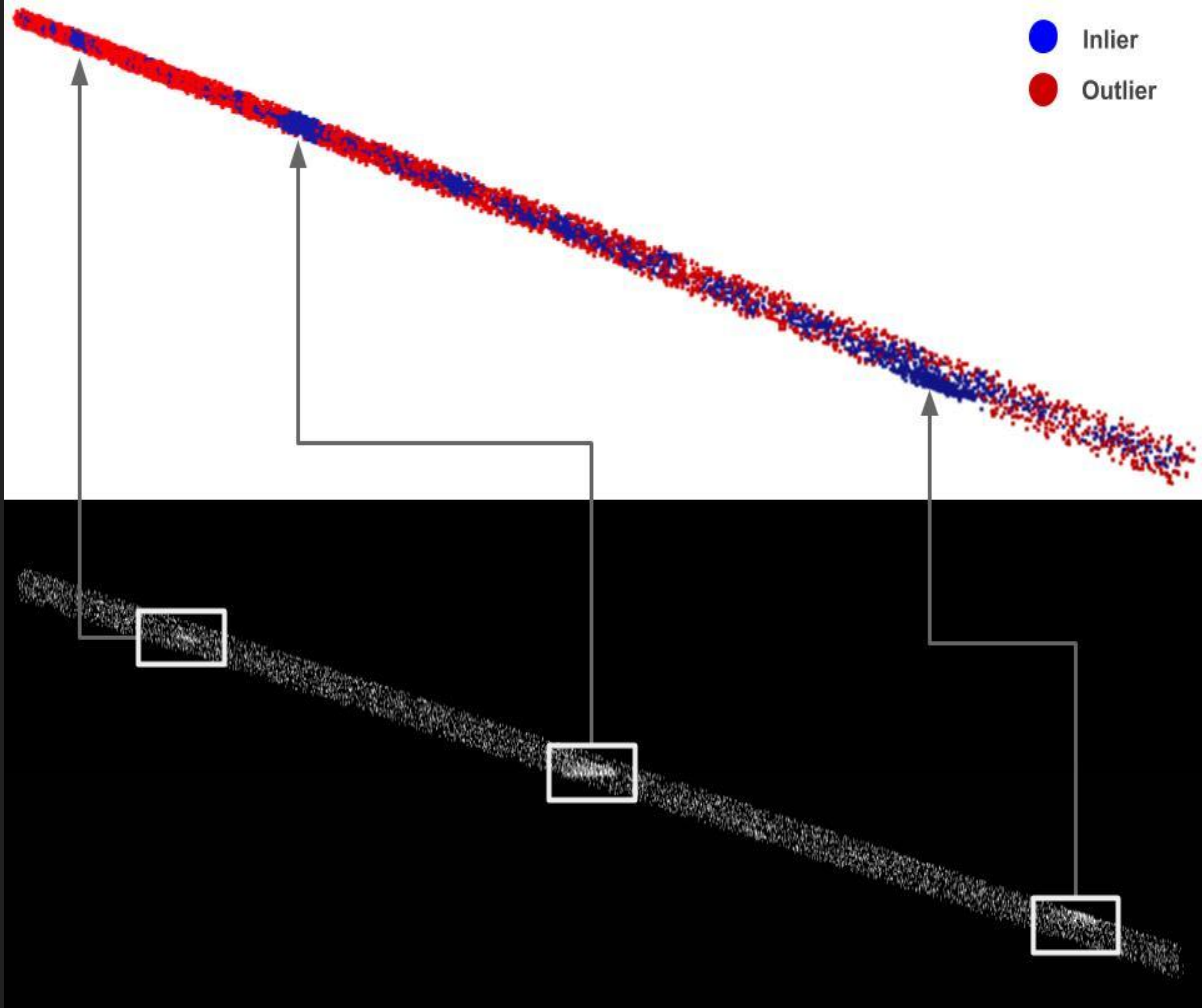


Proposed Solution: Radius-based Outlier Filter (RBOF)

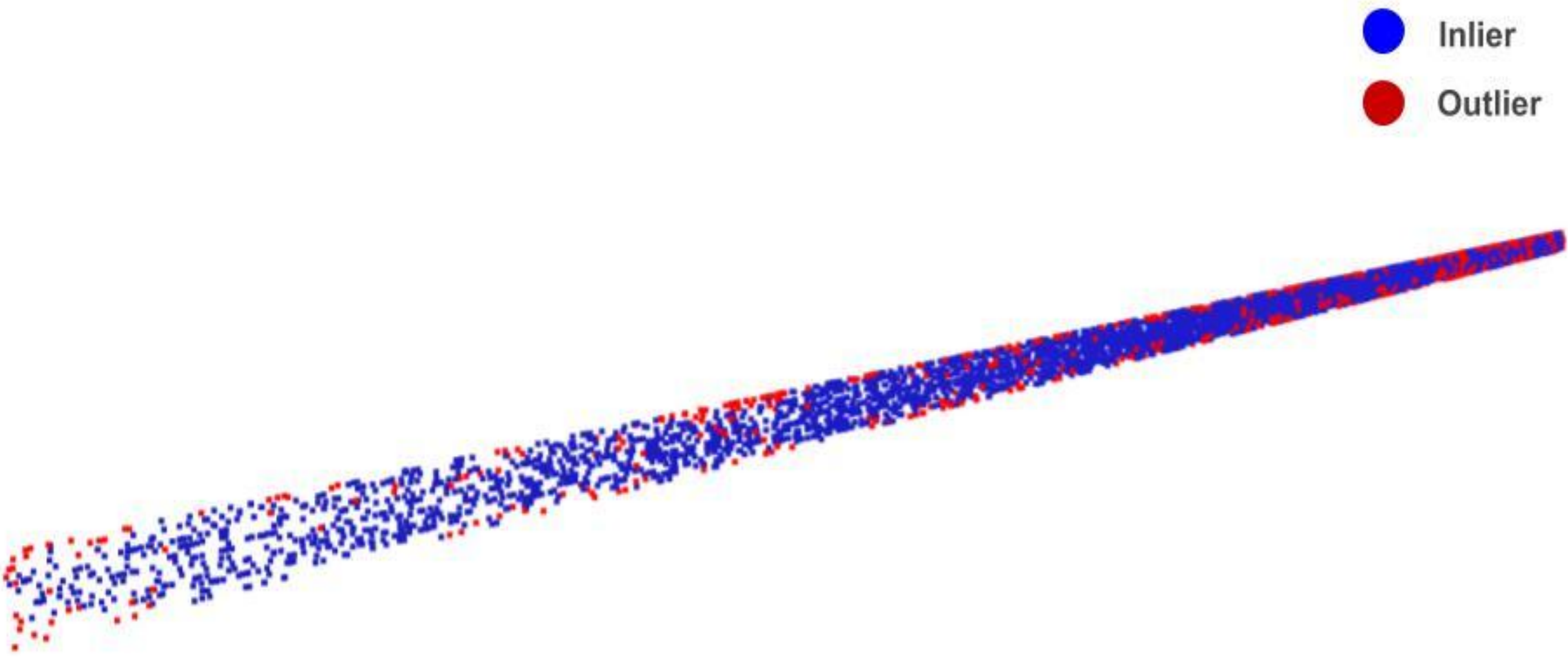
A point is an outlier if it contains <u>fewer</u> than specified neighbours



Outlier Detection Applied on An Event Timeslice



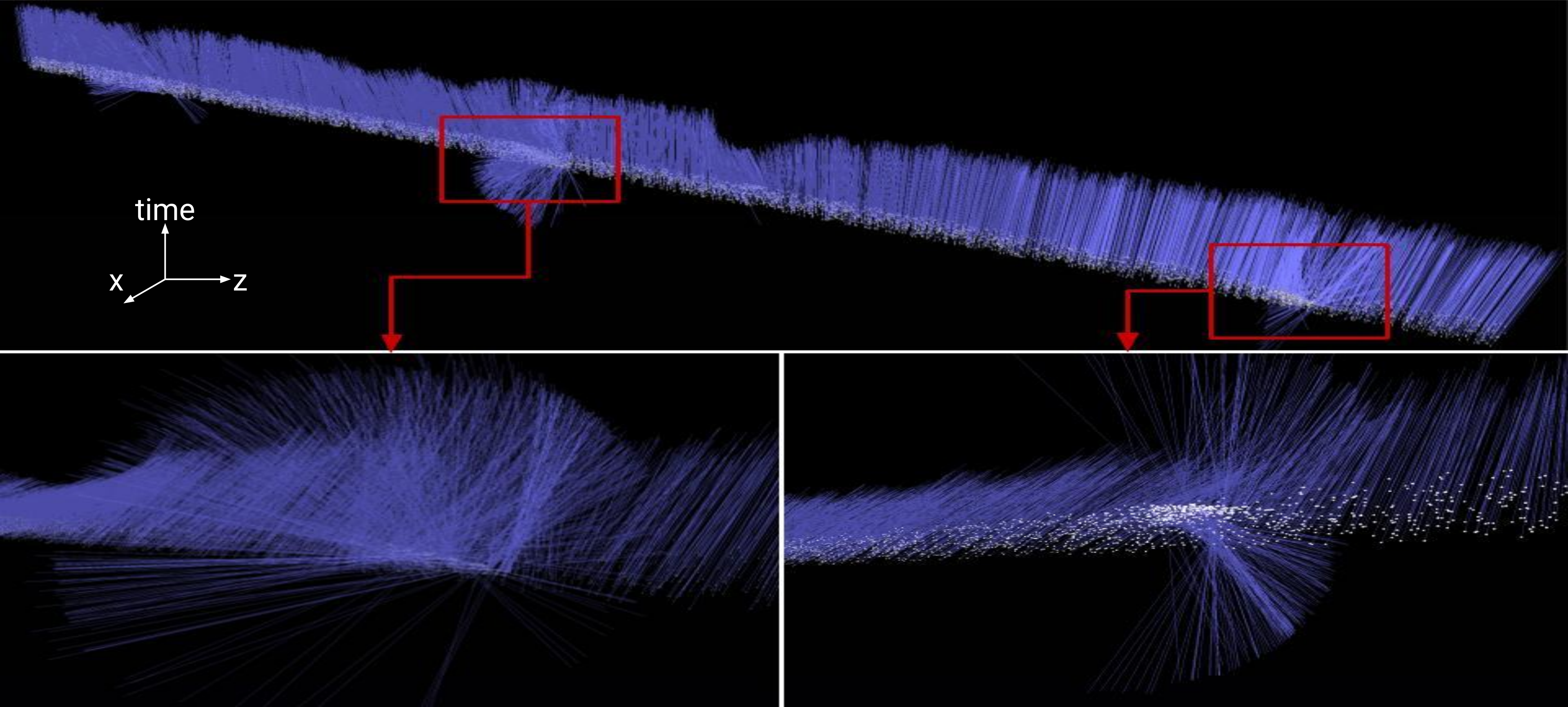
Outlier Detection Applied on Noise Timeslice



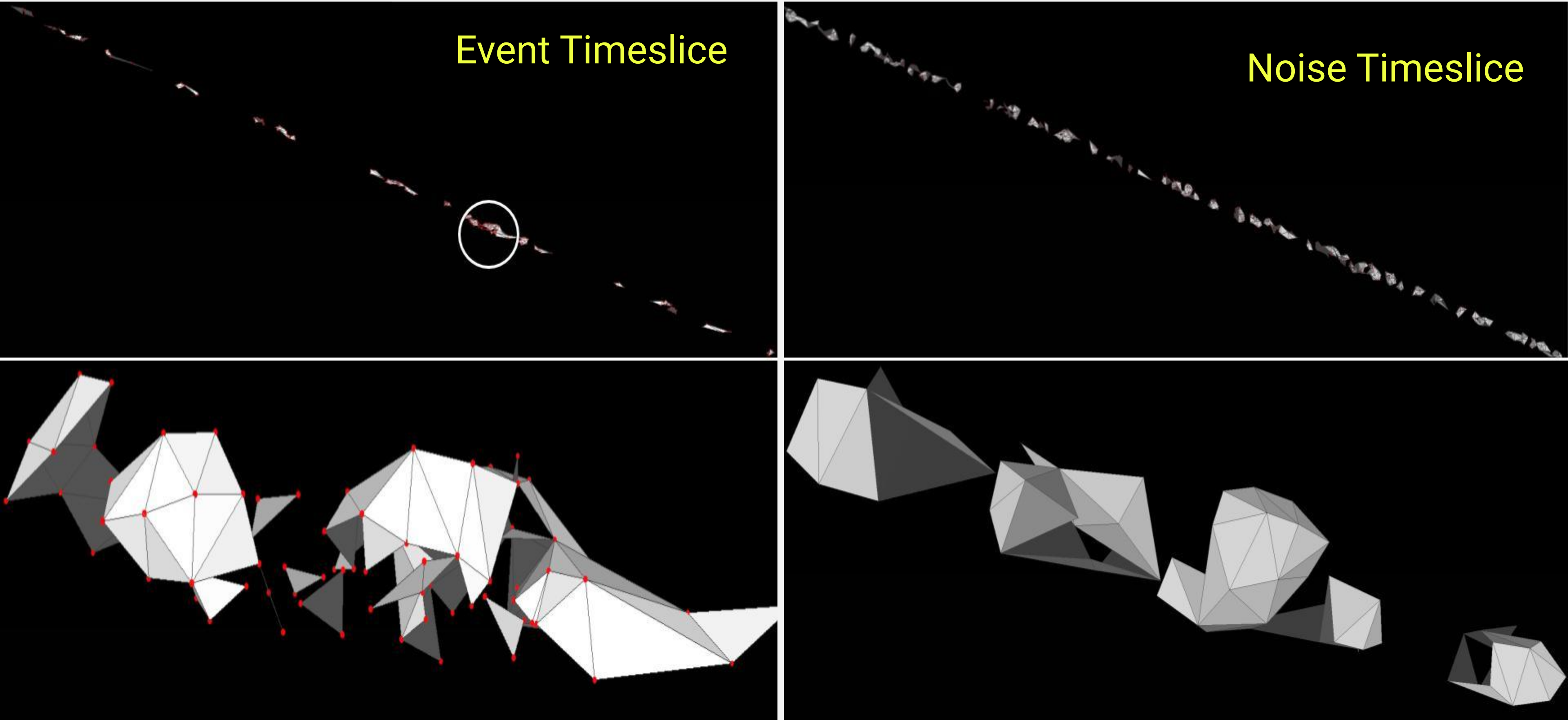
Timeslices After RBOF

Step 3: 3D Mesh Generation

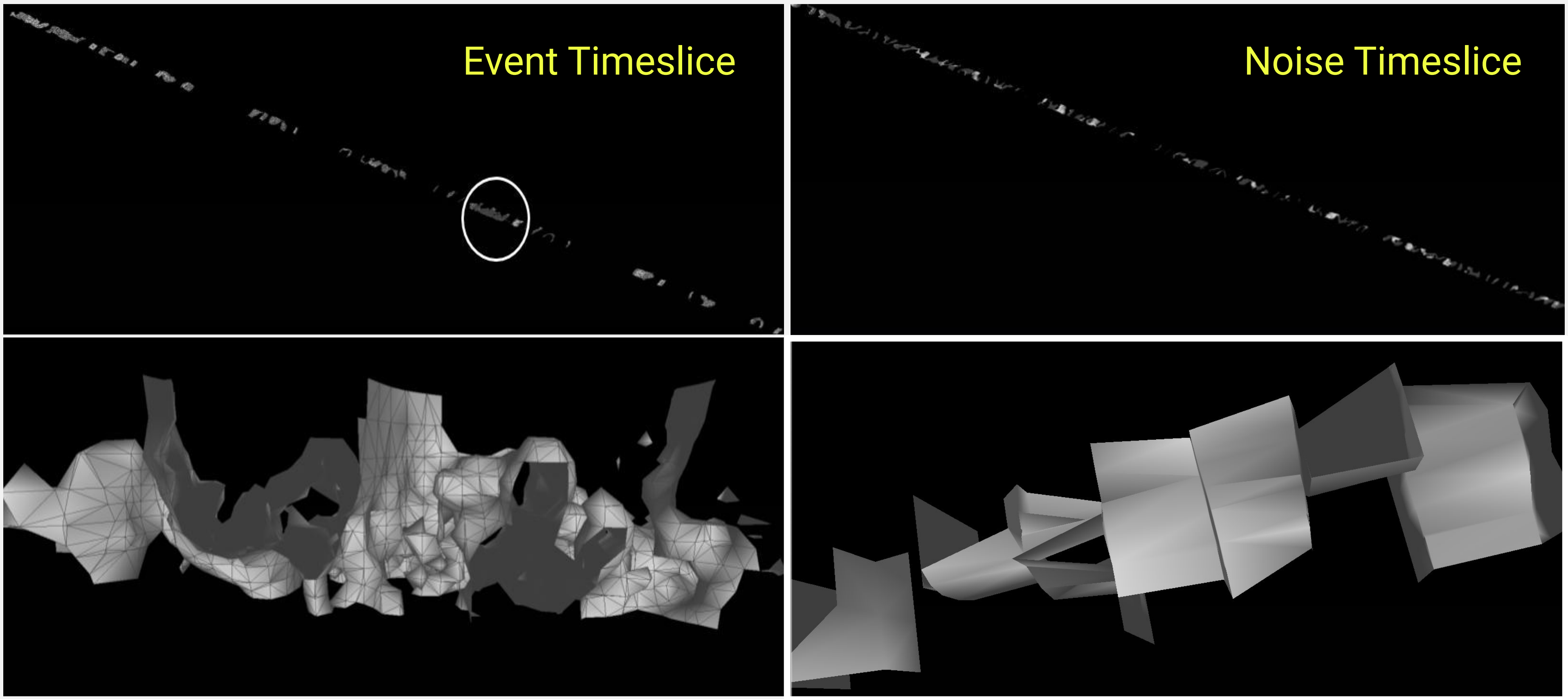
Computing Normals



Ball Pivoting Algorithm

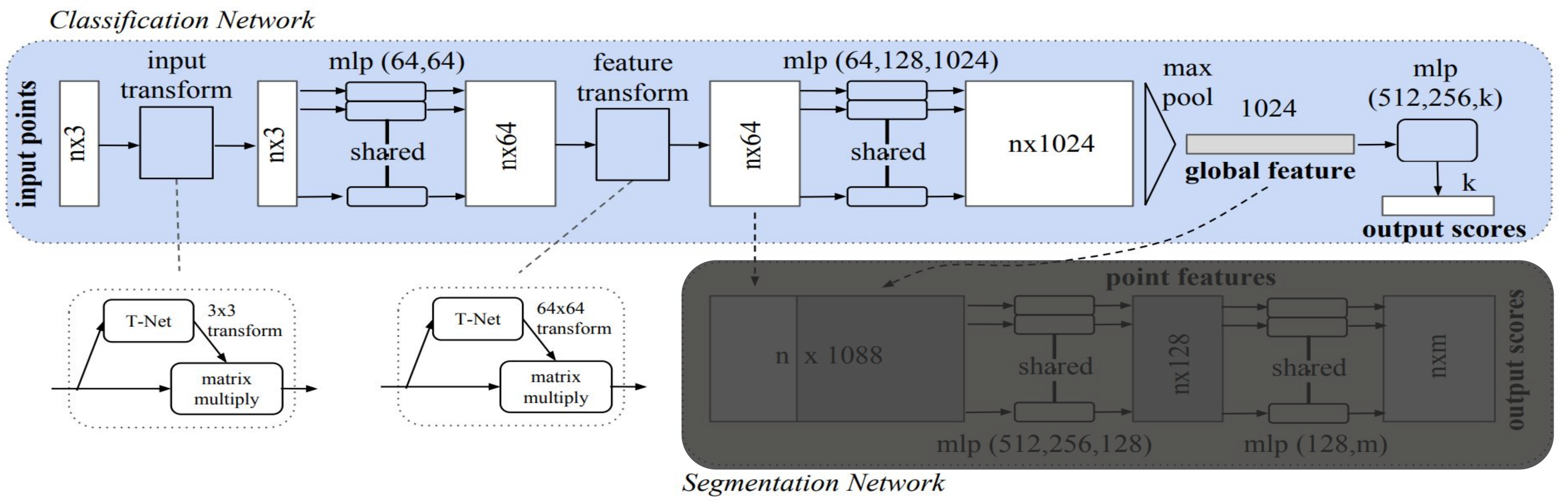


Surface Poisson Reconstruction



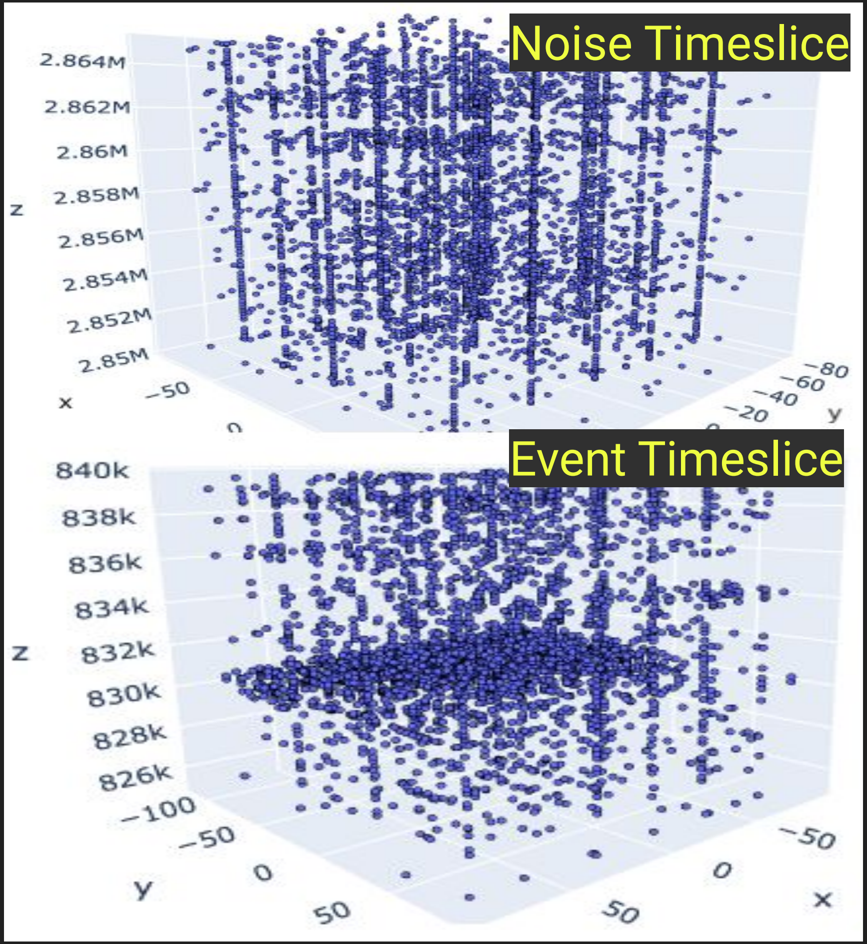
Step 4: Training & Evaluation

PointNet Background

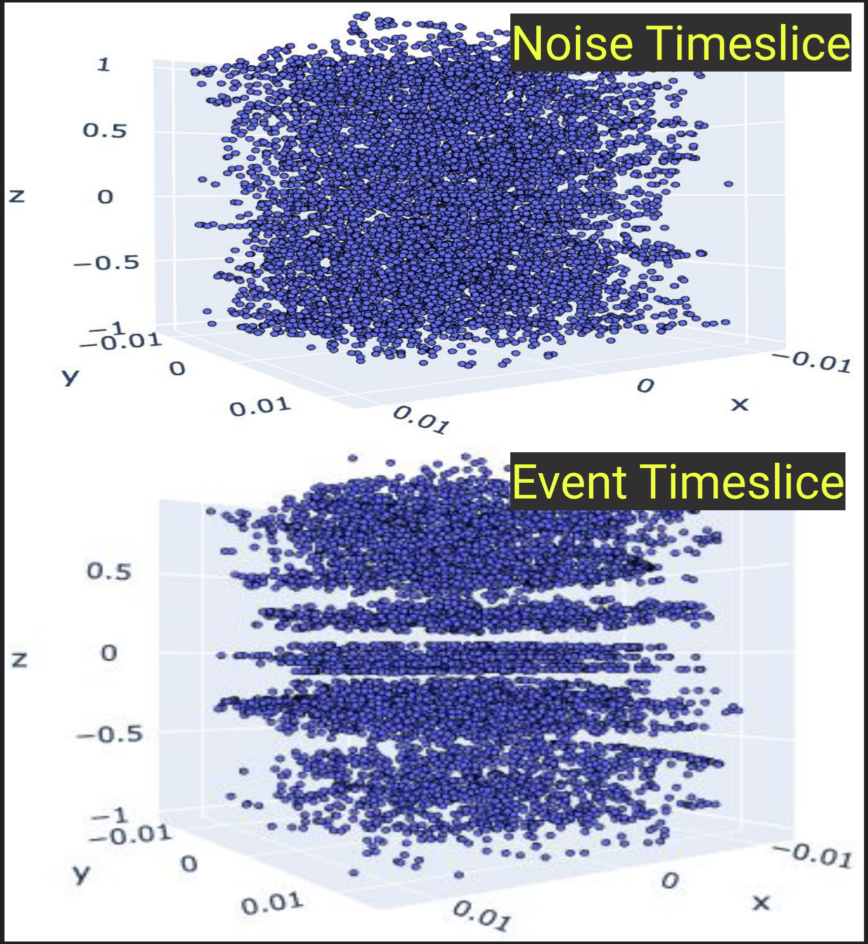


1. Fixed Sample of Points Per Cloud [7].
2. Unordered and invariant to permutations [7].
3. Invariant to transformations [7].

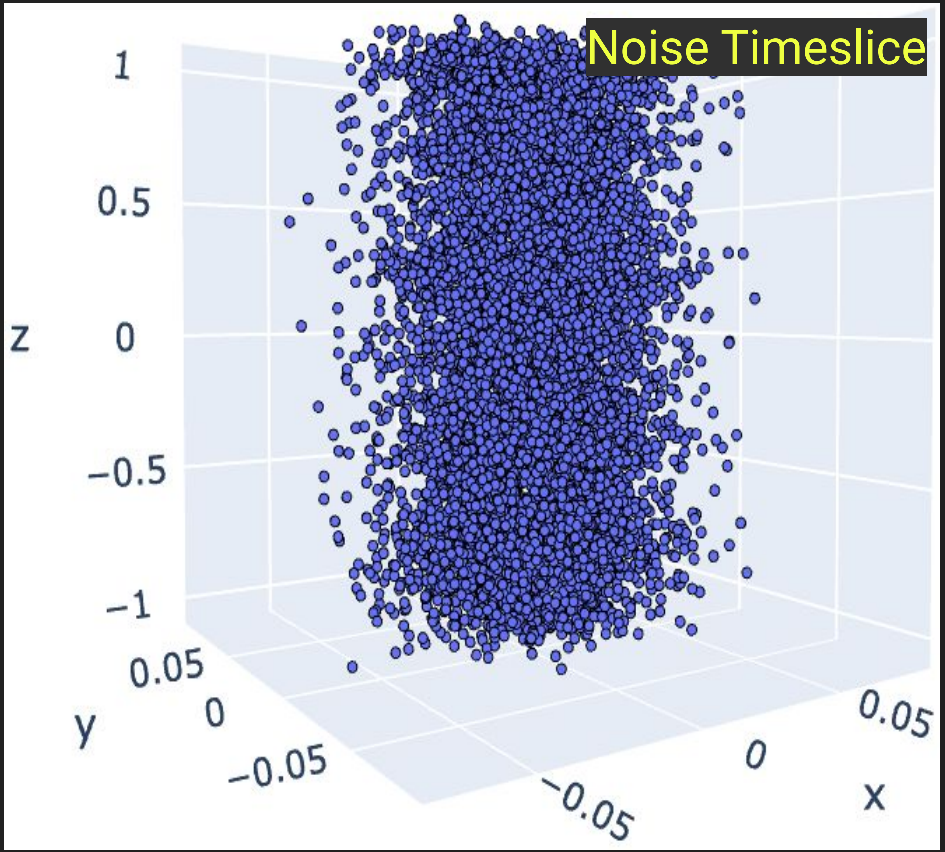
Sampling

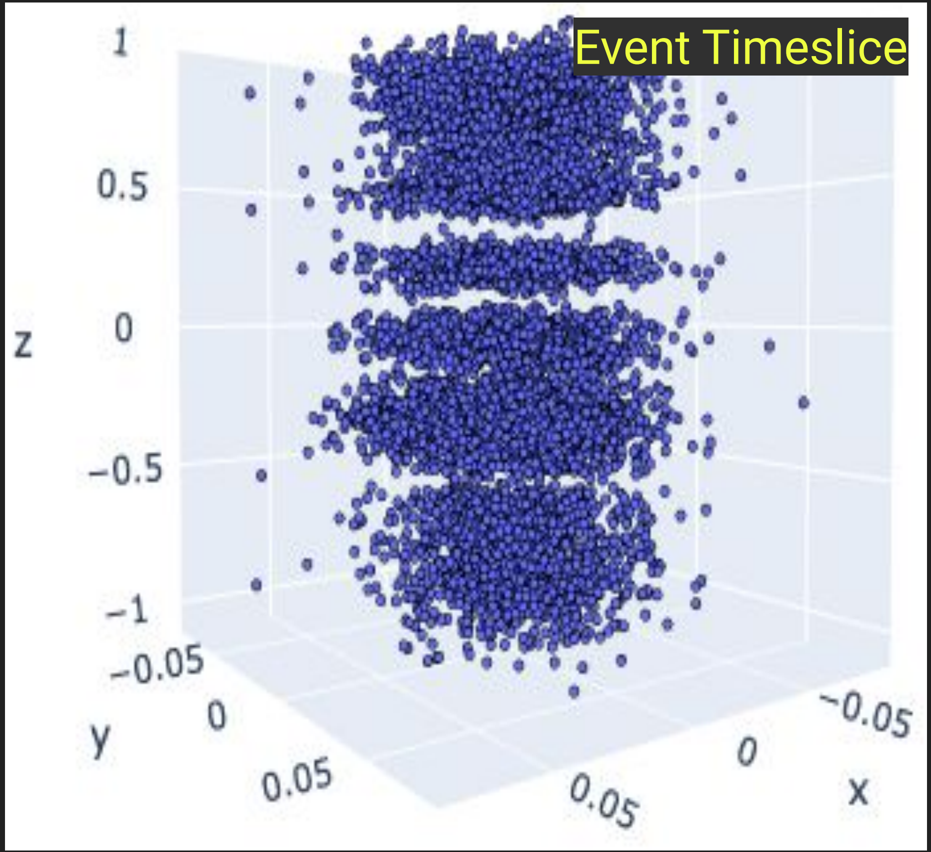


Normalisation



Jitter & Random Rotation





Evaluation Methodology

80/20 training-testing split

120 Epochs

Majority Voting Ensemble [11]

Results

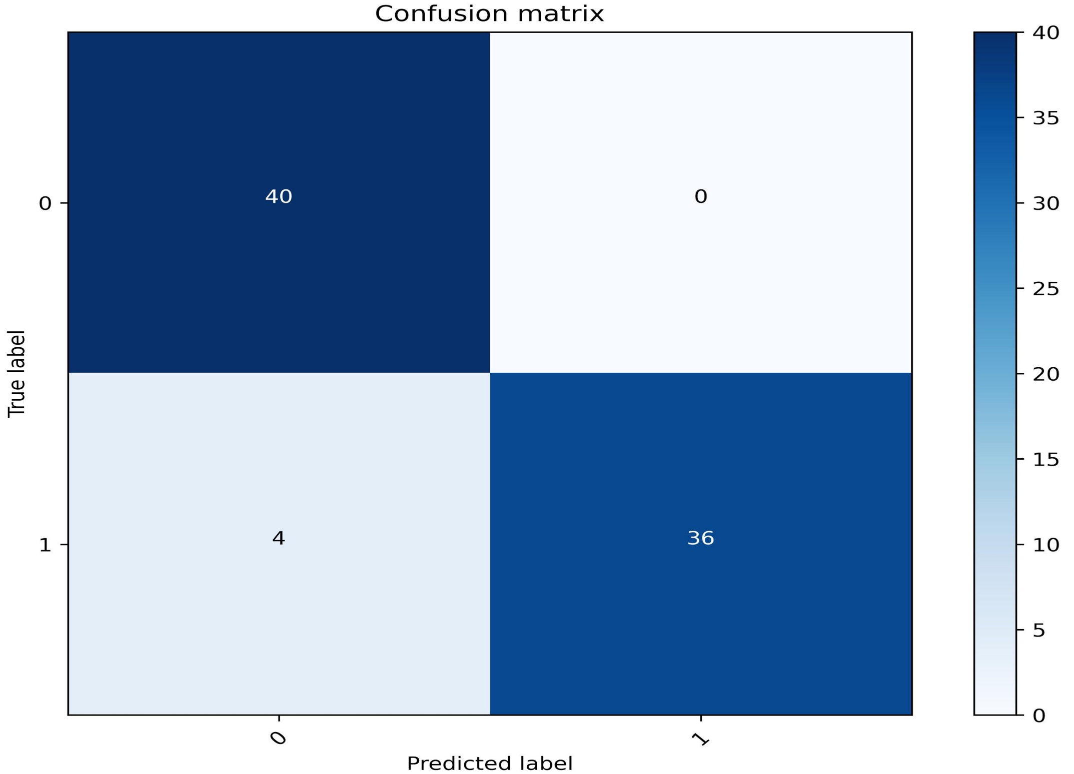
Accuracy

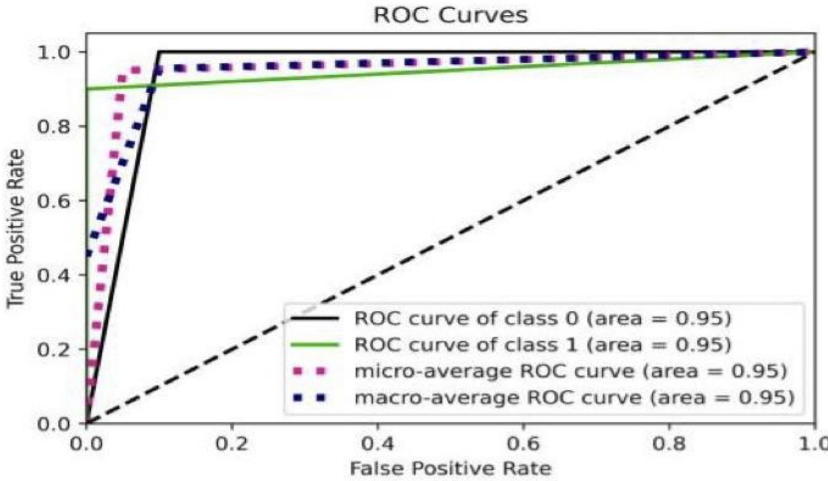
`x, y, time`	**95%** (loss = 0.003)
`x, z, time`	**90%** (loss = 0.006)
`y, z, time`	**99%** (loss = 0.005)
Ensemble 1: Hard Voting	**97%**
Ensemble 2: Soft Voting	**90%**

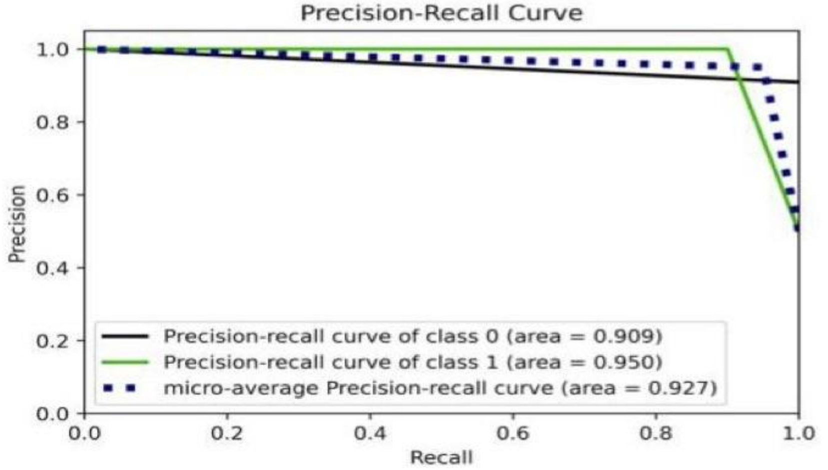
Table 7.1: Accuracy Scores

Dataset 1: `x y time`

	precision	recall	F1-score	support
class_0	0.91	1.00	0.95	40
class_1	1.00	$\boxed{0.90}$	0.95	40

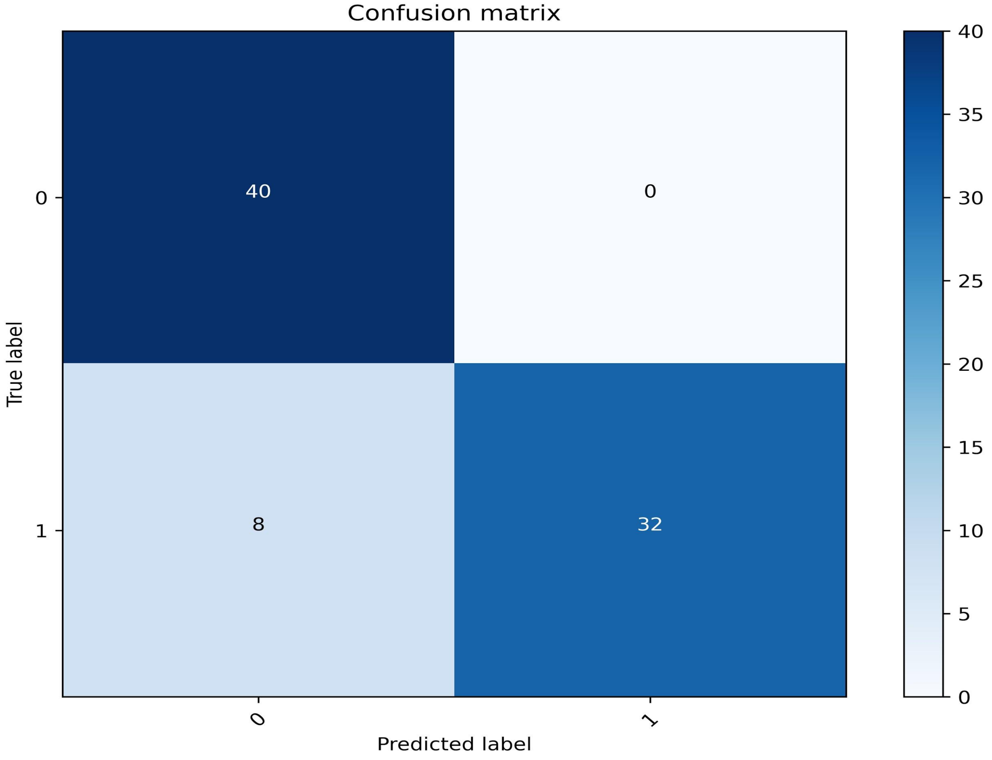


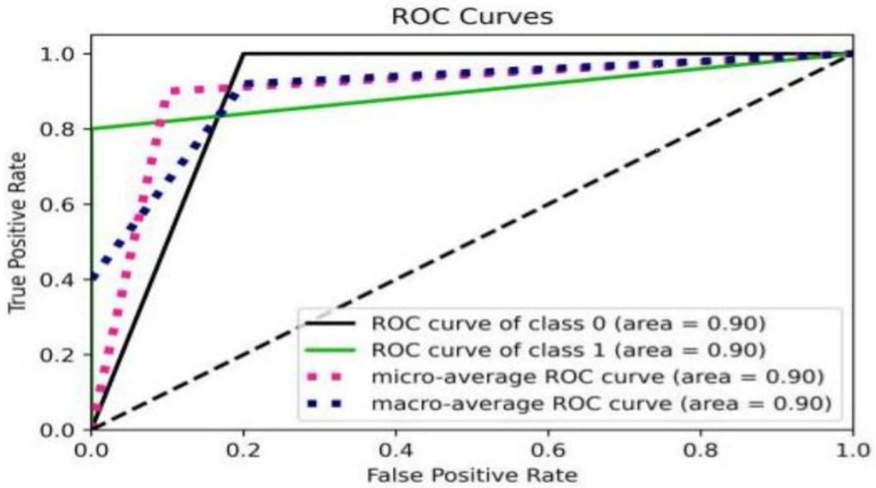


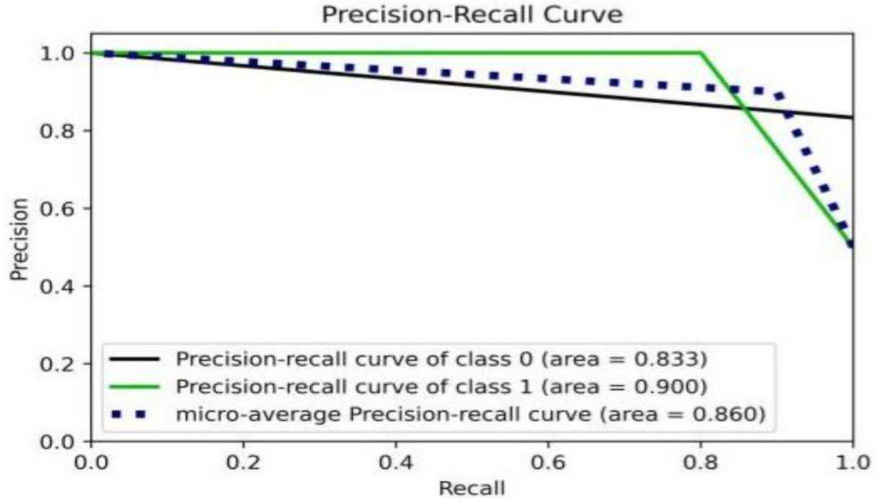


Dataset 2: `x z time`

	precision	recall	F1-score	support
class_0	0.83	1.00	0.91	40
class_1	1.00	$\boxed{0.80}$	0.89	40

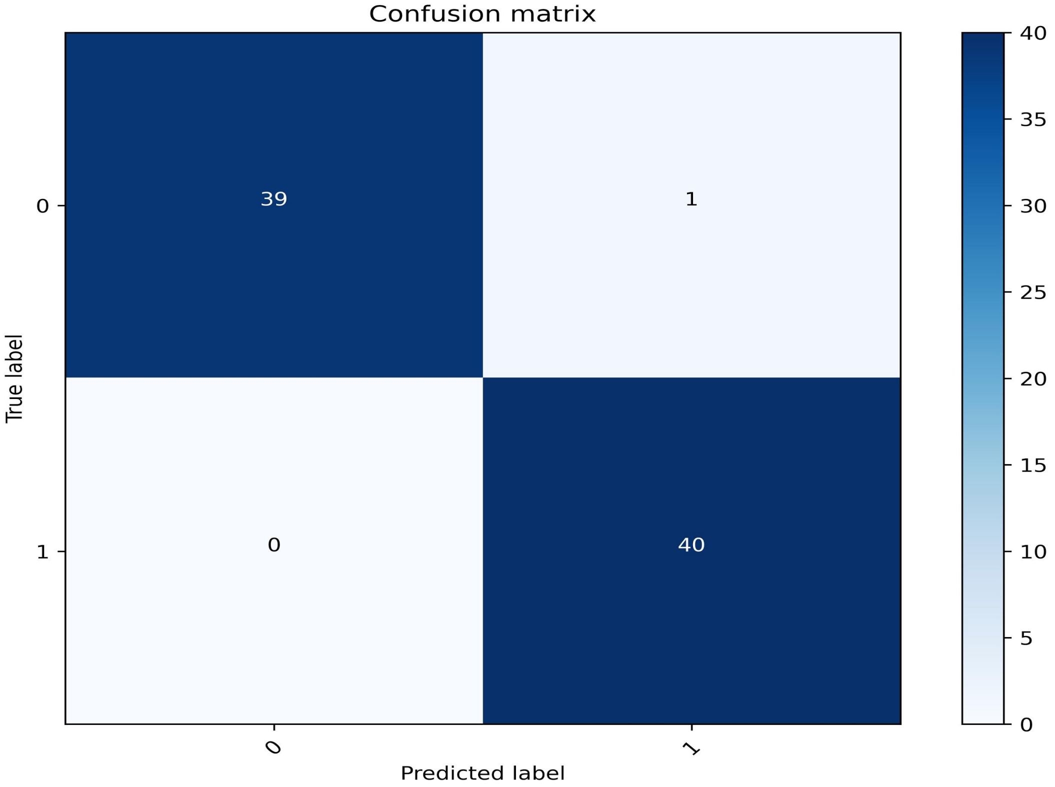


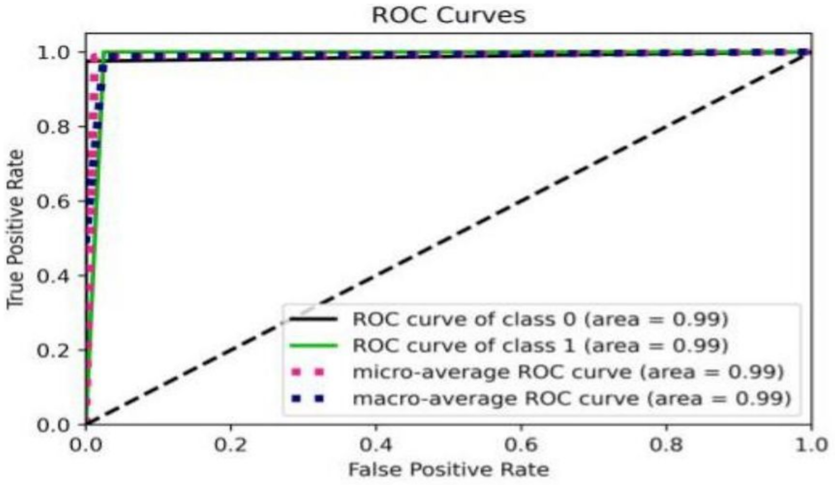


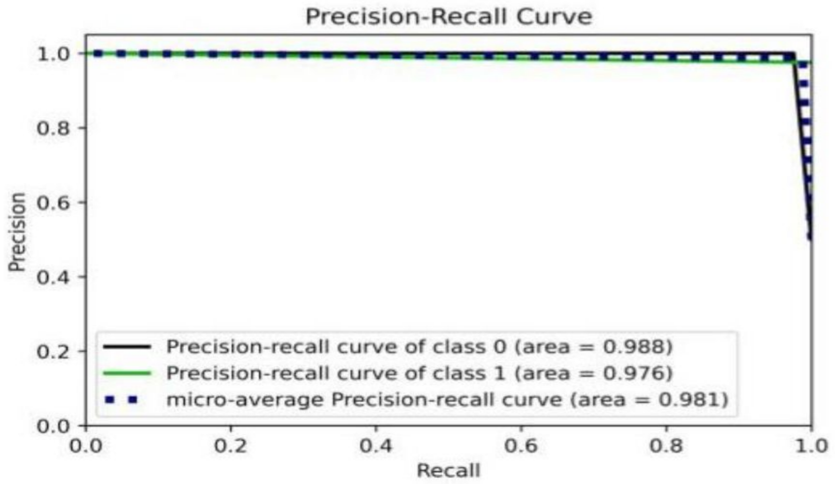


Dataset 3: `y z time`

	precision	recall	F1-score	support
class_0	1.00	0.97	0.99	40
class_1	0.98	$\boxed{1.00}$	0.99	40

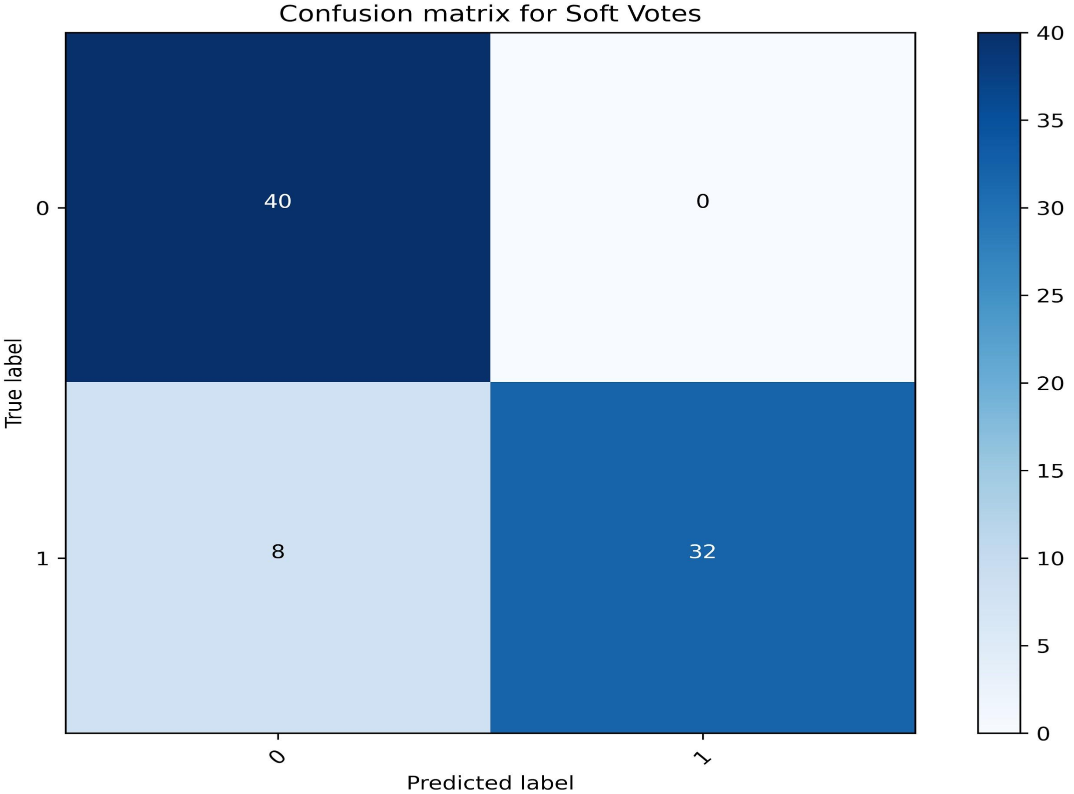


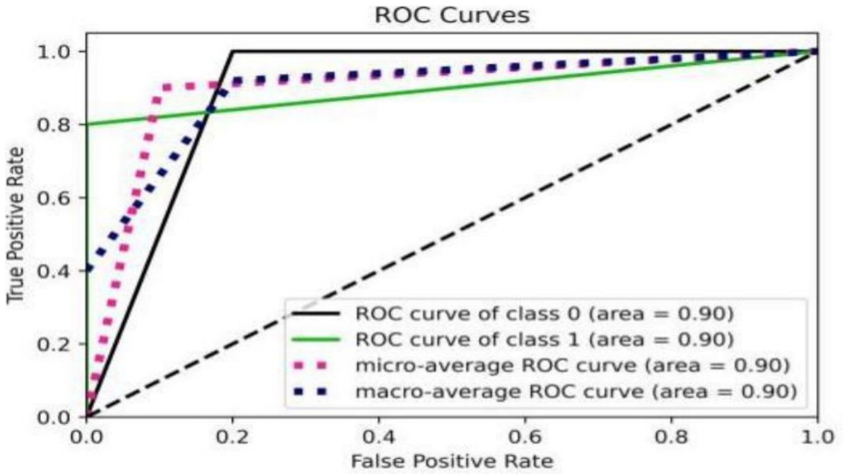


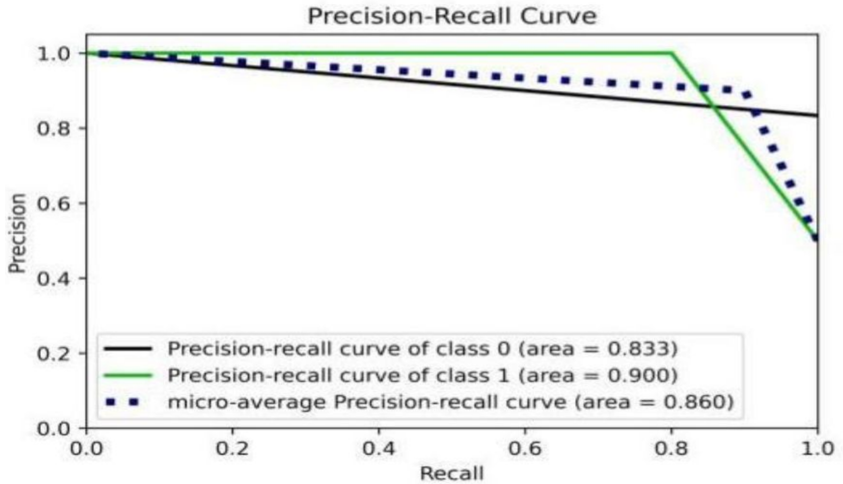


Ensemble 1: Soft Voting

	precision	recall	f1-score	support
class_0	0.93	1.00	0.91	40
class_1	1.00	$\boxed{0.80}$	0.89	40

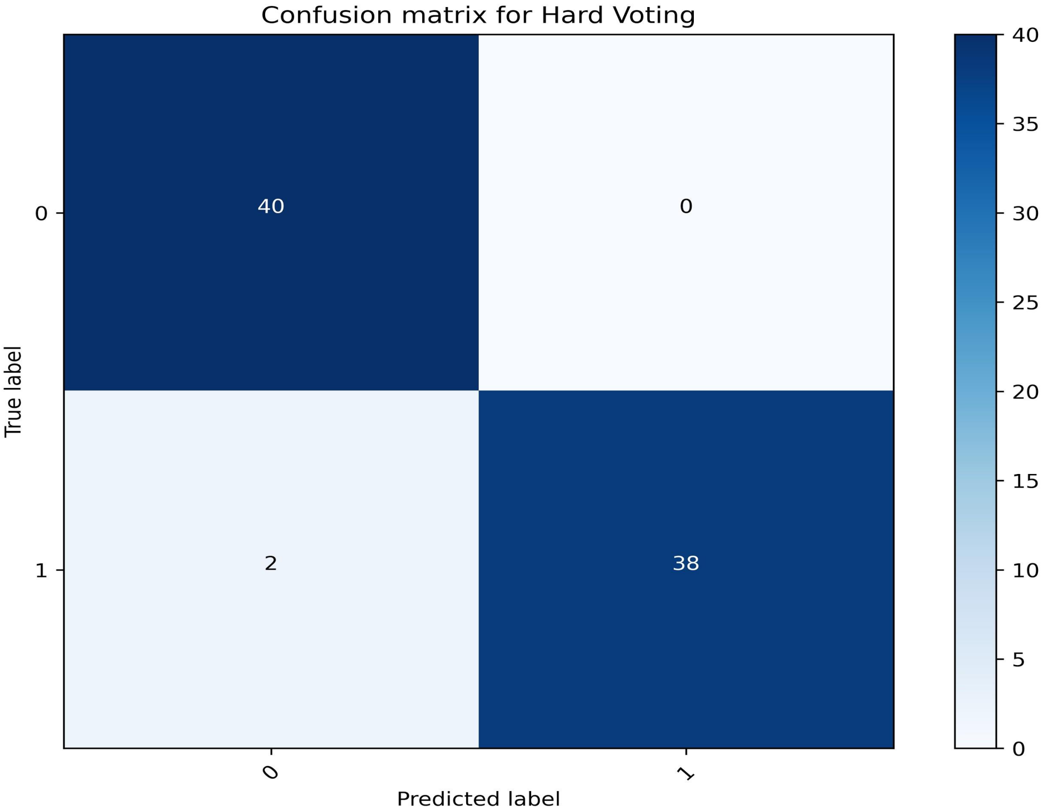


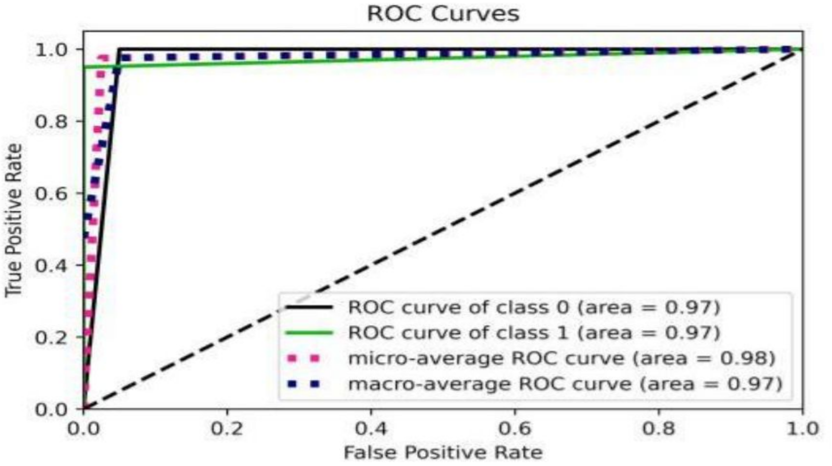


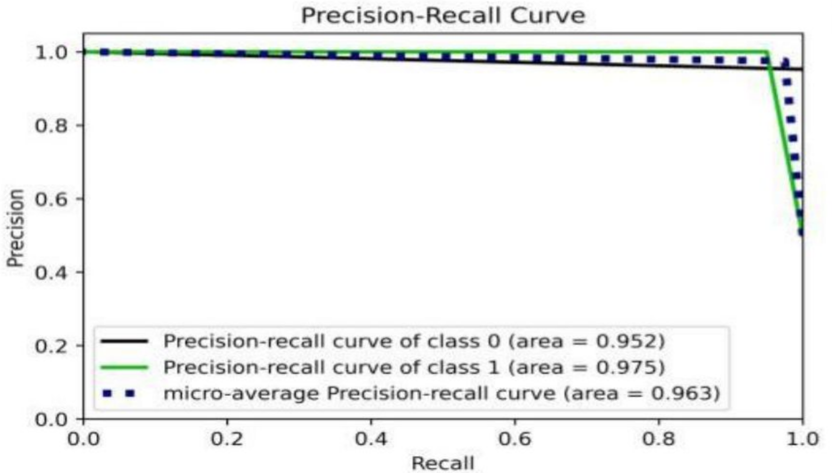


Ensemble 2: Hard Voting

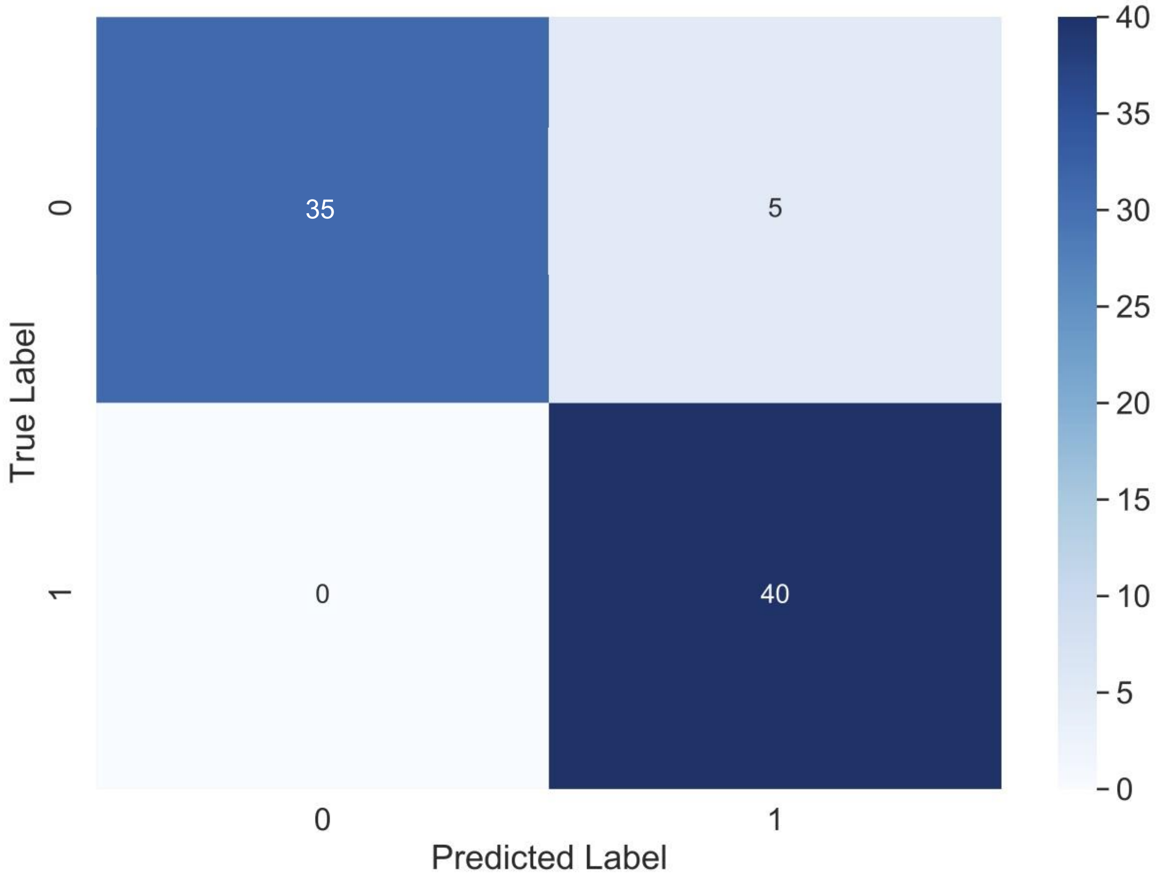
	precision	recall	f1-score	support
class_0	0.95	1.00	0.98	40
class_1	1.00	$\boxed{0.95}$	0.97	40







L1 Trigger Confusion Matrix



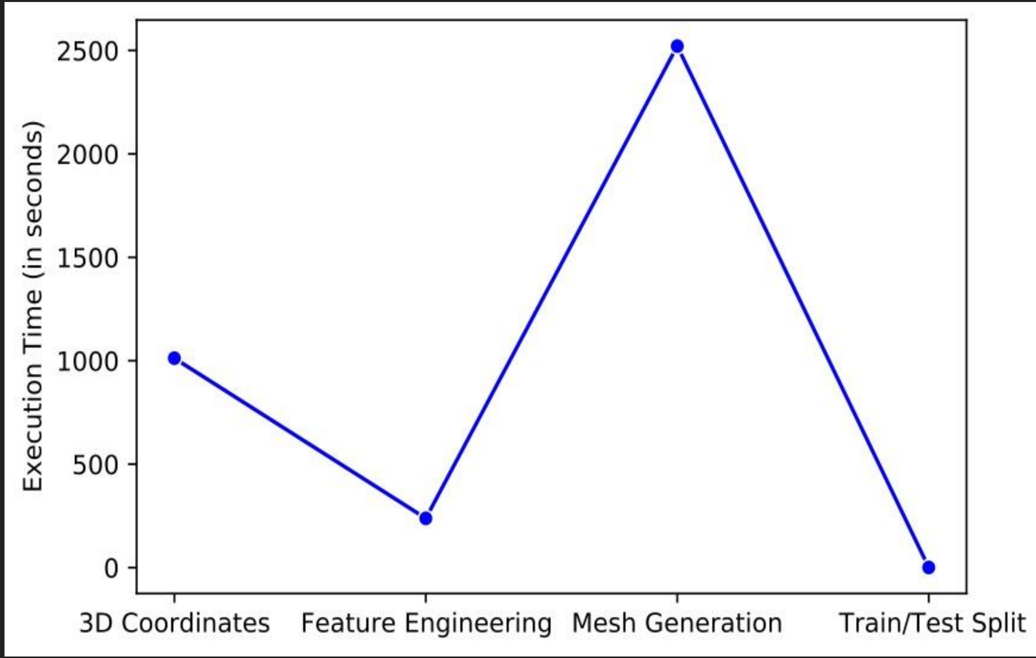
Additional Metrics

Pre-Processing Execution Time
~ 1.04 hours

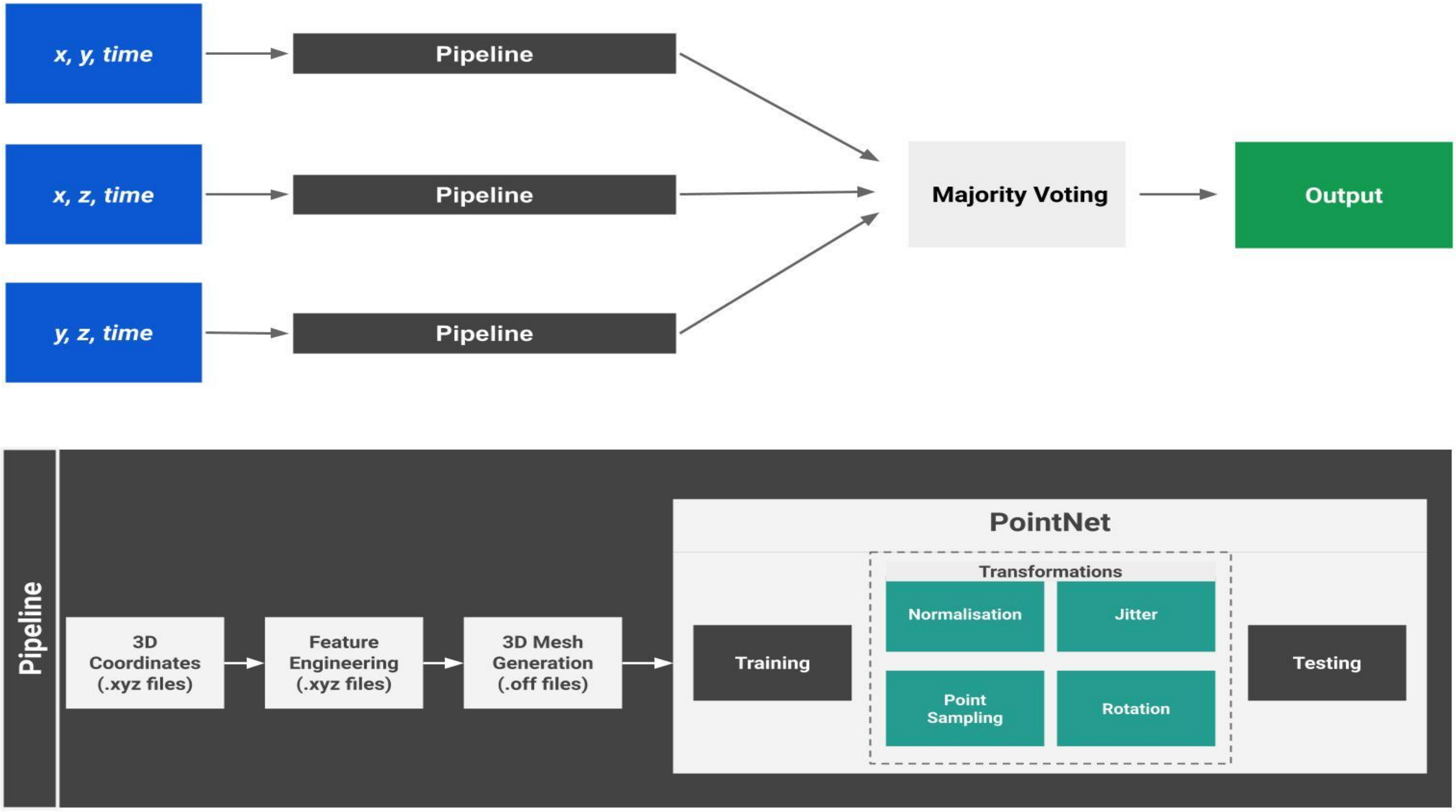
PointNet Execution Time
~ 3.2 hours

PointNet Throughput
0.34 instances / second

PointNet Energy [12]:
62 g of CO2 for 120 epochs

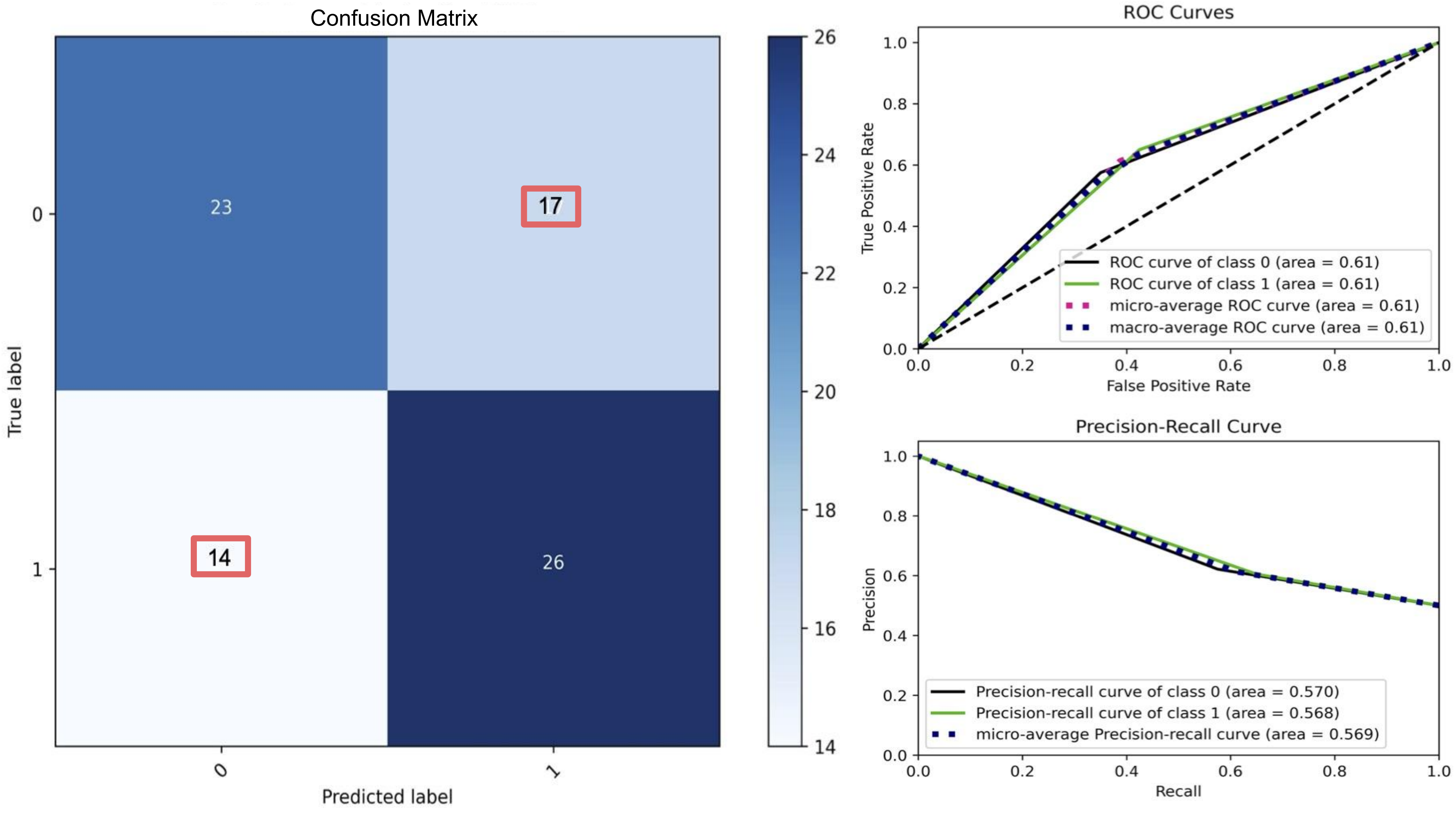


Pre-Processing Pipeline Execution Time for Each Stage



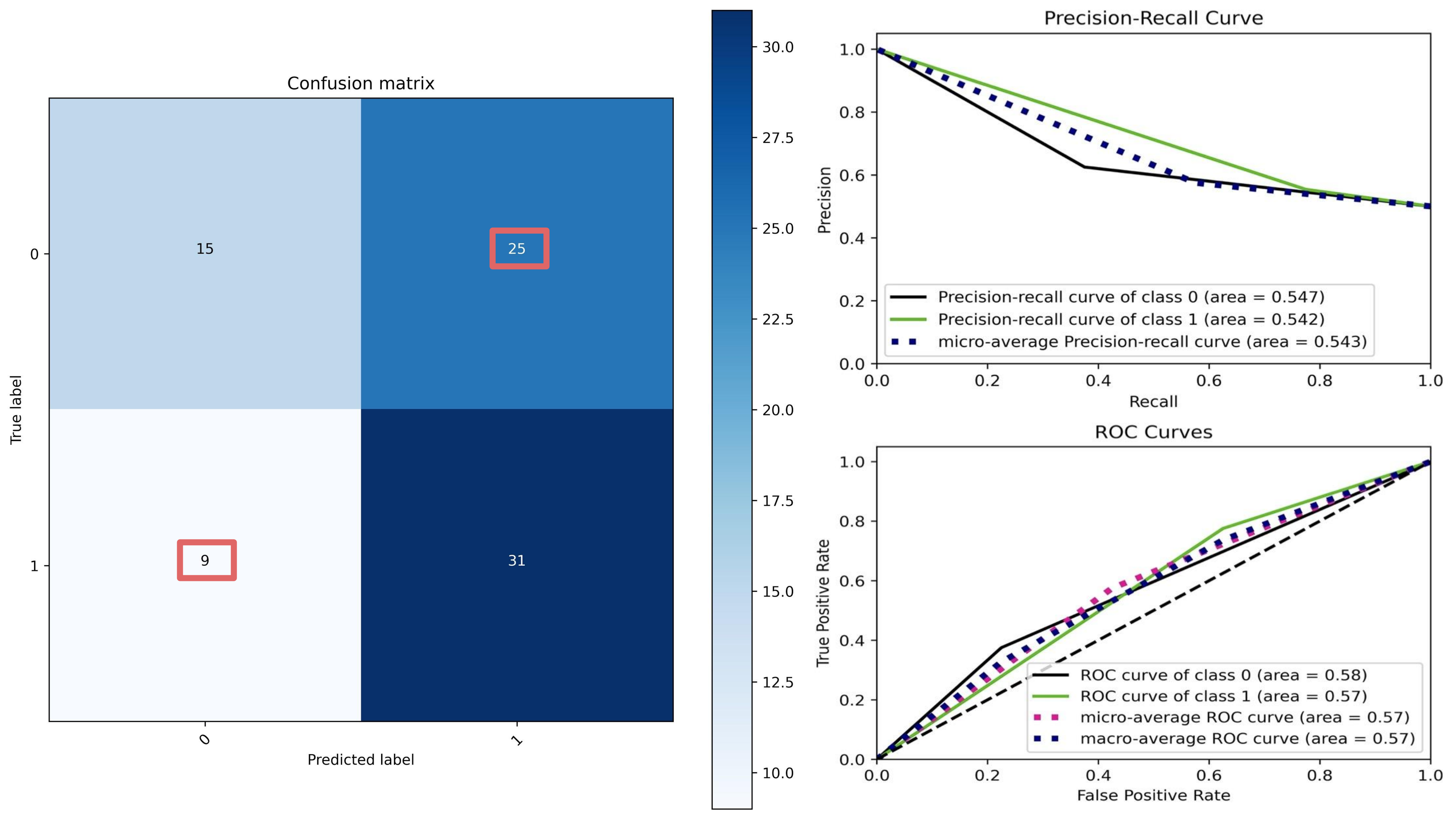
x, y, time
x, z, time
y, z, time
Pipeline
Pipeline
Pipeline
Majority Voting
Output
Pipeline
3D Coordinates (.xyz files)
Feature Engineering (.xyz files)
3D Mesh Generation (.off files)
PointNet
Training
Transformations
Normalisation
Jitter
Point Sampling
Rotation
Testing

3D PointNet



Confusion Matrix
23
17
14
26
True label
0
1
Predicted label
0
1
26
24
22
20
18
16
14
ROC Curves
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve of class 0 (area = 0.61)
ROC curve of class 1 (area = 0.61)
micro-average ROC curve (area = 0.61)
macro-average ROC curve (area = 0.61)
Precision-Recall Curve
Precision
Recall
Precision-recall curve of class 0 (area = 0.570)
Precision-recall curve of class 1 (area = 0.568)
micro-average Precision-recall curve (area = 0.569)

4D PointNet


Confusion matrix
15
25
9
31
0
1
True label
Predicted label
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
10.0
Precision-Recall Curve
Precision
Recall
Precision-recall curve of class 0 (area = 0.547)
Precision-recall curve of class 1 (area = 0.542)
micro-average Precision-recall curve (area = 0.543)
ROC Curves
True Positive Rate
False Positive Rate
ROC curve of class 0 (area = 0.58)
ROC curve of class 1 (area = 0.57)
micro-average ROC curve (area = 0.57)
macro-average ROC curve (area = 0.57)

Energy Inference

PointNet $\neq$ Regression

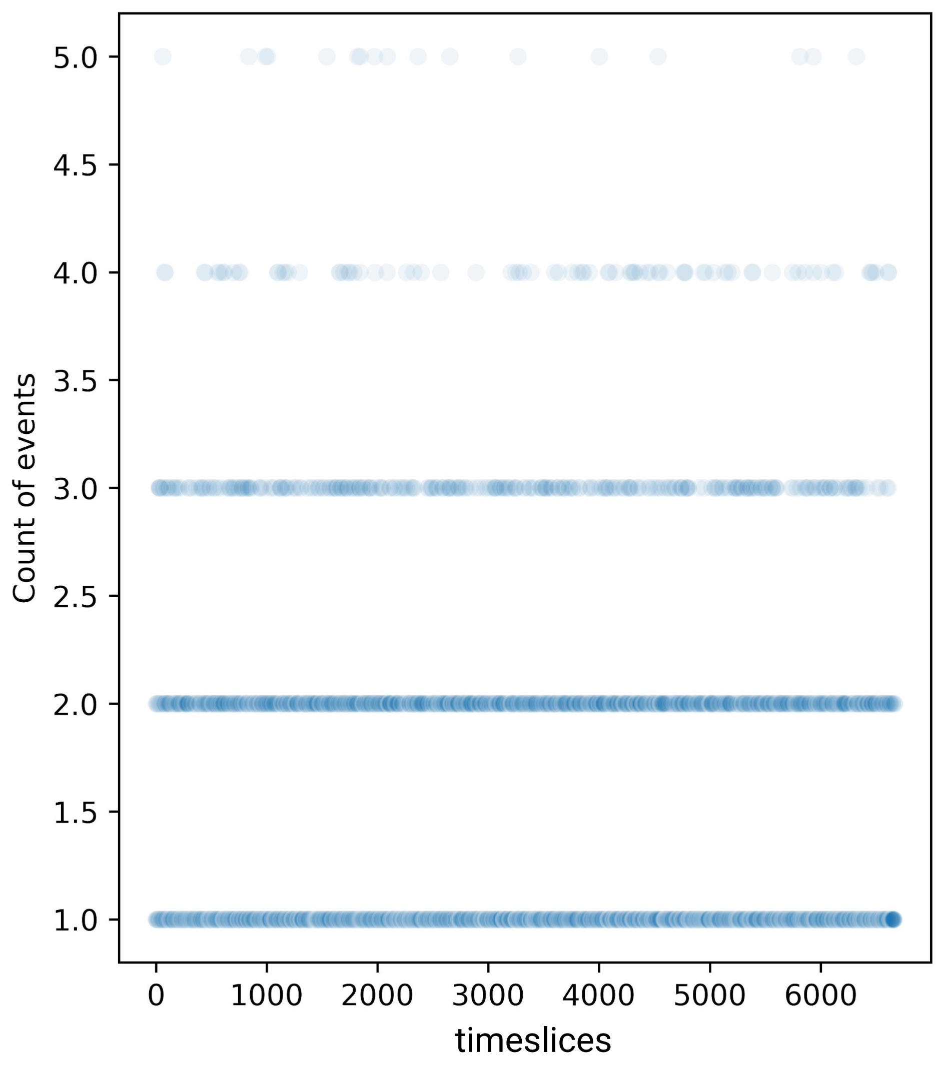
Data

Key Variables:

`x y z`
`time`
`energy`
`group`

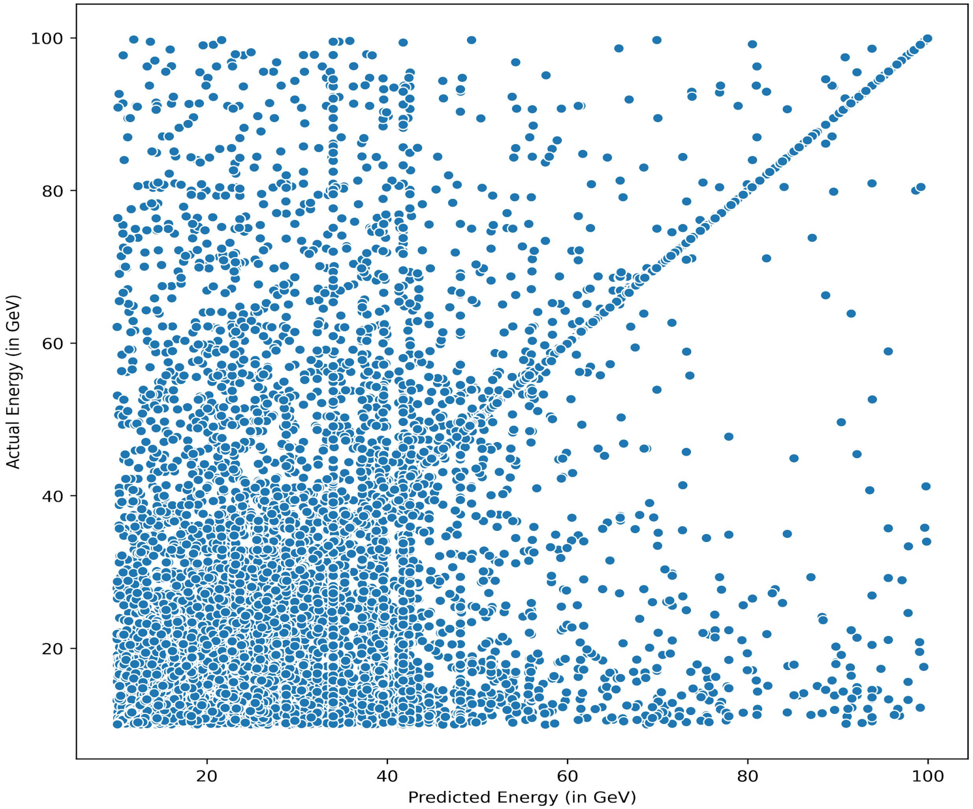
No significant correlations

66/34 split & **50** evaluation samples

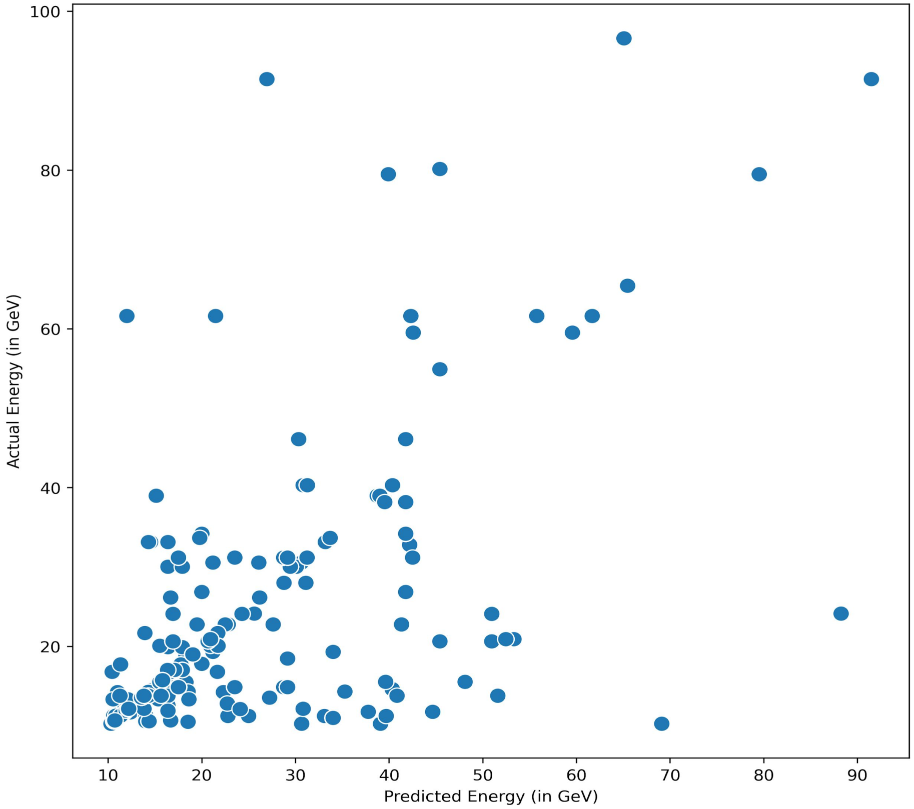


Decision Trees

	Training	Testing	Additional
MSE	100.3	116.3	269.0
R2	0.83	$\boxed{0.80}$	$\boxed{0.31}$



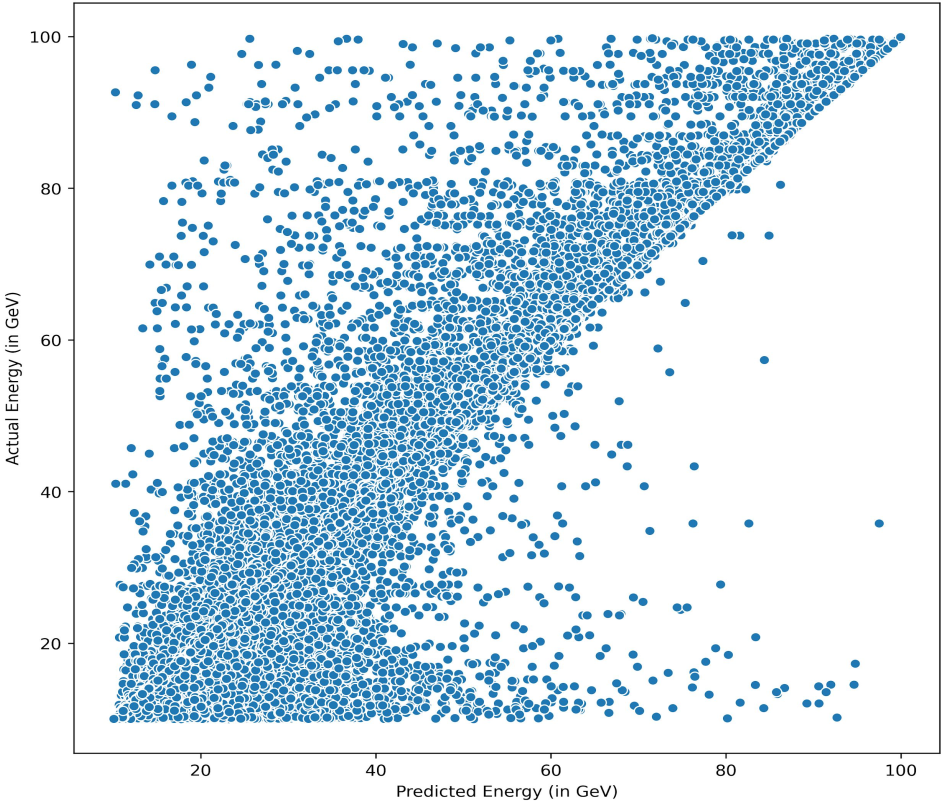
Testing Data



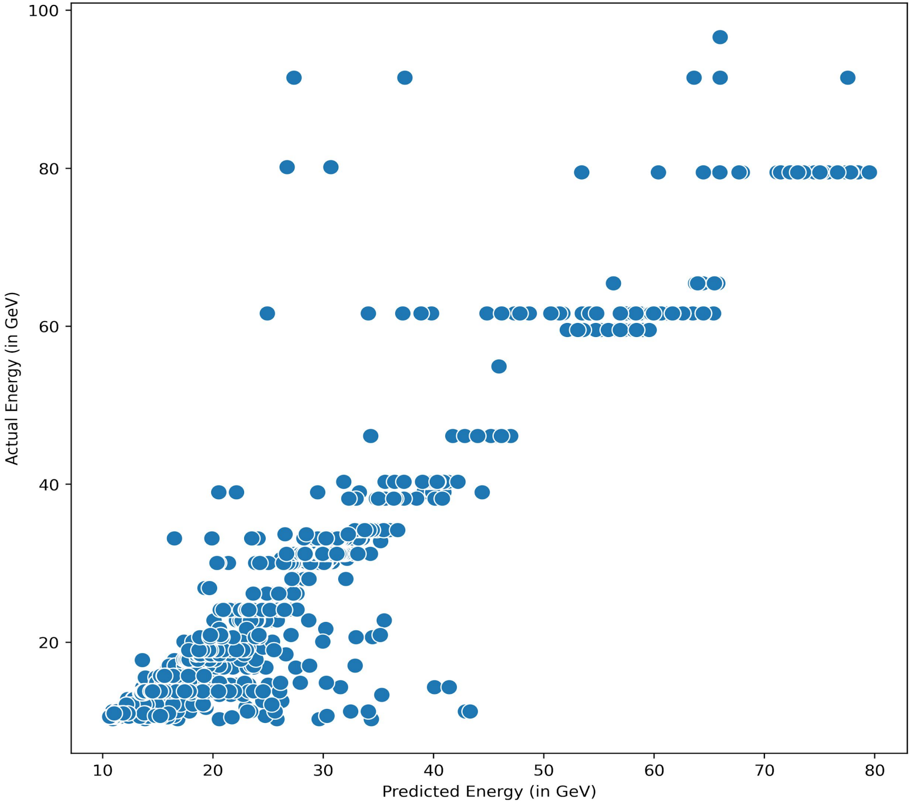
Additional Data

Random Forest Bootstrapping

	Training	Testing	Additional
MSE	2.6	15.93	79.8
R2	0.99	0.97	0.80



Testing Data



Additional Data

Limitations & Recommendations

Feature Engineering → PointNet 2

Limited Testing (80 examples) → Expand Test Cases

Surface Reconstruction Algorithm → PowerCrust

PointNet unsuitable for regression	→	Faster RCNN
Energy results analysis	→	Address uncertainties
~3 minutes for classifying new timeslice	→	GPU and Parallelisation

Revisiting Research Questions

RQ 2.0	Can the KM3NeT dataset be effectively represented using 3D meshes?
RQ 2.1	Which meshing algorithm would be suitable for representing data?

RQ 1.0 Can PointNet classify noise timeslices and event timeslices?

RQ 1.1 Can PointNet obtain a Recall score of 0.9 for the positive class?

RQ 3.0

Can PointNet be extended for energy inference from events?

Thank You!

References

[1] Serguei Chatrchyan, V Khachatryan, AM Sirunyan, W Adam, B Arnold, H Bergauer, T Bergauer, M Dragicevic, M Eichberger, J Ero, et al. 2010. Performance of the CMS Level-1 trigger during commissioning with cosmic ray muons and LHC beams. Journal of Instrumentation 5 (2010)
[2] Halina Abramowicz, Allen Caldwell, and Ralph Sinkus. 1995. Neural network based electron identification in the ZEUS calorimeter. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 365, 2-3 (1995), 508–517.
[3] P Adamson, C Andreopoulos, KE Arms, R Armstrong, DJ Auty, S Avvakumov, DS Ayres, B Baller, B Barish, PD Barnes Jr, et al. 2008. Study of muon neutrino disappearance using the Fermilab Main Injector neutrino beam. Physical Review D 77, 7 (2008), 072002.
[4] Adrian-Martinez, S., Ageron, M., Aharonian, F., Aiello, S., Albert, A., Ameli, F., ... & Anton, G. (2016). Letter of intent for KM3NeT 2.0. Journal of Physics G: Nuclear and Particle Physics, 43(8), 084001.
[5] Karaś, K. (2019). Data processing pipeline for the KM3NeT neutrino telescope (Doctoral dissertation, Universiteit van Amsterdam).
[6] Aiello, S., Albert, A., Garre, S. A., Aly, Z., Ameli, F., Andre, M., ... & Ardid, M. (2020). Event reconstruction for KM3NeT/ORCA using convolutional neural networks. arXiv preprint arXiv:2004.08254.
[7] Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 652-660).
[8] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1, No. 2). Cambridge: MIT press.
[9] Zubair, S., Yan, F., & Wang, W. (2013). Dictionary learning based sparse coefficients for audio classification with max and average pooling. Digital Signal Processing, 23(3), 960-970.[10] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Desmaison, A. (2019). Pytorch: An imperative style, high-performance deep learning library. In Advances in neural information processing systems (pp. 8026-8037).
[11] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.
[12] Anthony, L. F. W., Kanding, B., & Selvan, R. (2020). Carbontracker: Tracking and predicting the carbon footprint of training deep learning models. arXiv preprint arXiv:2007.03051.
[13] Sakia, R. M. (1992). The Box-Cox transformation technique: a review. Journal of the Royal Statistical Society: Series D (The Statistician), 41(2), 169-178.

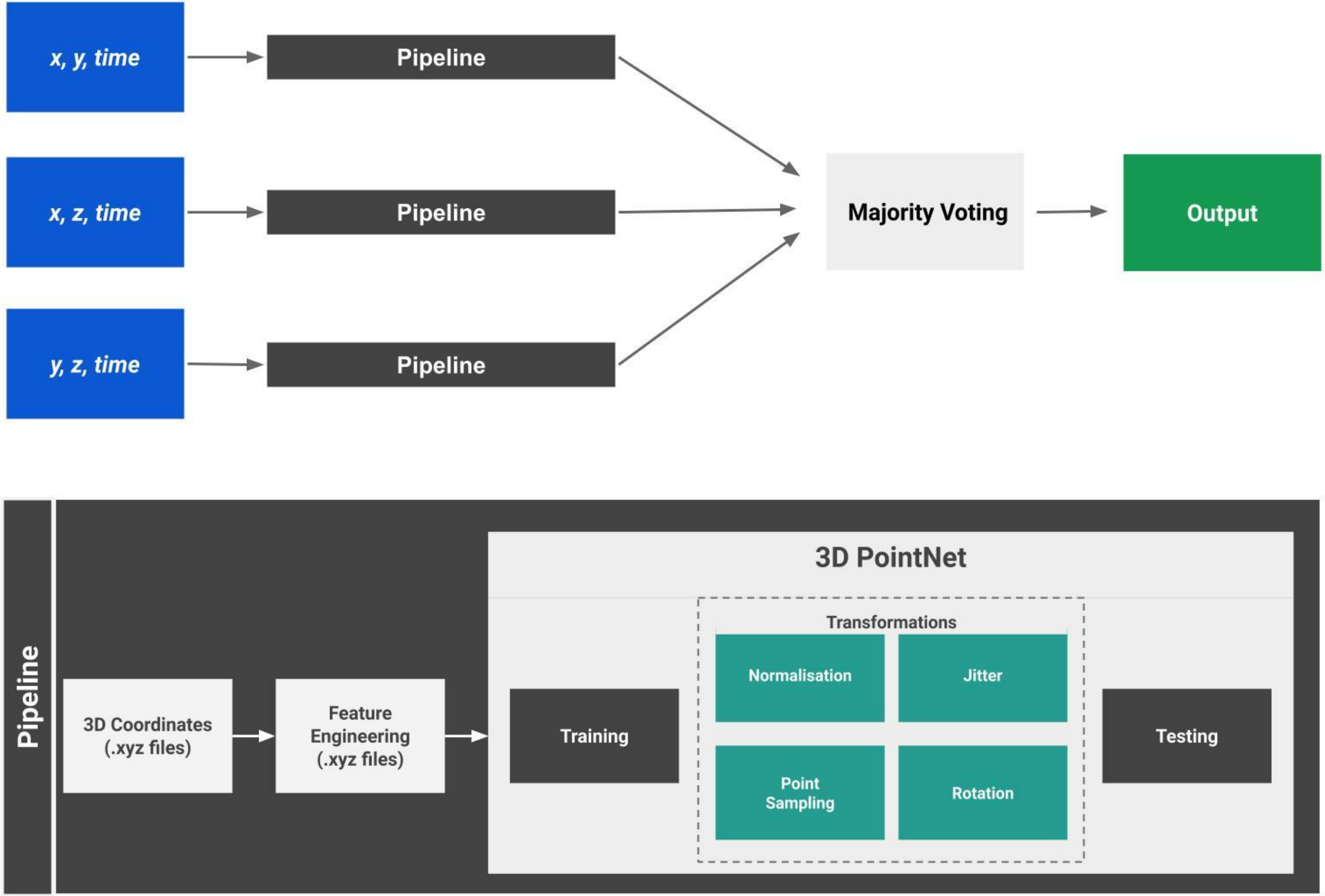
Appendix

Modifications

1D Convolutions [8]

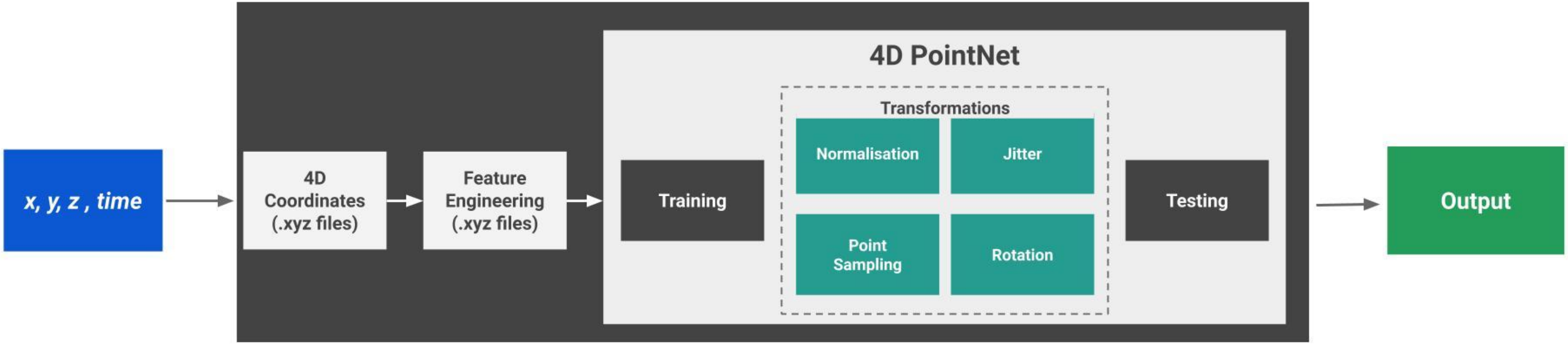
Average Max Pooling instead of Global Max Pooling [8, 9]

Negative Log Likelihood Loss instead of Cross Entropy Loss [8, 10]


x, y, time
x, z, time
y, z, time
Pipeline
Pipeline
Pipeline
Majority Voting
Output
Pipeline
3D Coordinates (.xyz files)
Feature Engineering (.xyz files)
3D PointNet
Transformations
Normalisation
Jitter
Point Sampling
Rotation
Training
Testing

3D PointNet

4D PointNet



x, y, z , time
4D Coordinates (.xyz files)
Feature Engineering (.xyz files)
4D PointNet
Training
Transformations
Normalisation
Jitter
Point Sampling
Rotation
Testing
Output