Exploring DeepSet Neural Network for Modeling Neutron Multi-Scatters

Adeep Sen

30 June 2026

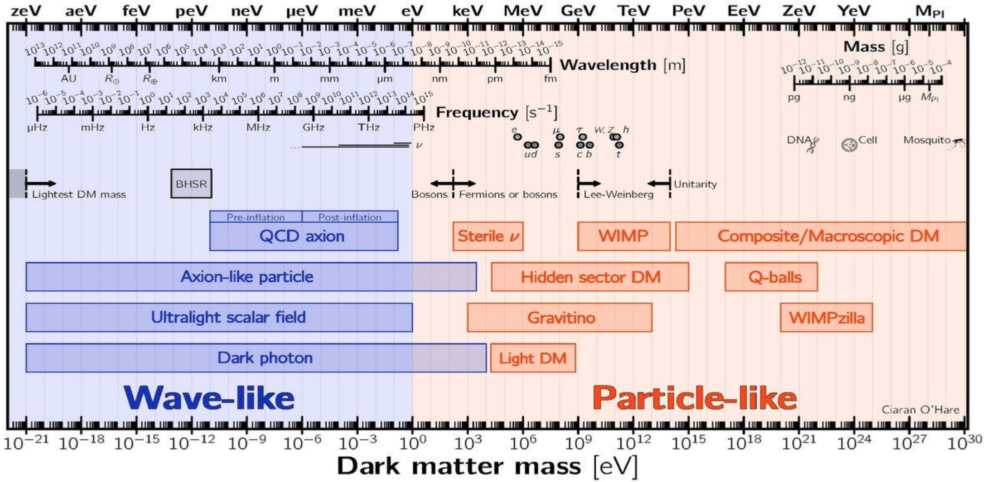
Columbia University in the City of New York

XENON

What is dark matter?

- Mysterious invisible substance making up 27% of the universe
- Holds galaxies together gravitationally
- Causes gravitational lensing

- What is it though?
  - Nobody knows



Dark matter candidates. From O'Hare (2024), Wikimedia Commons, CC BY 4.0.

The XENON Dark Matter Detector

- Searching for Weakly Interacting Massive Particles (WIMPs)
- Liquid Xenon Dual Phase Time Projection chamber (TPC)

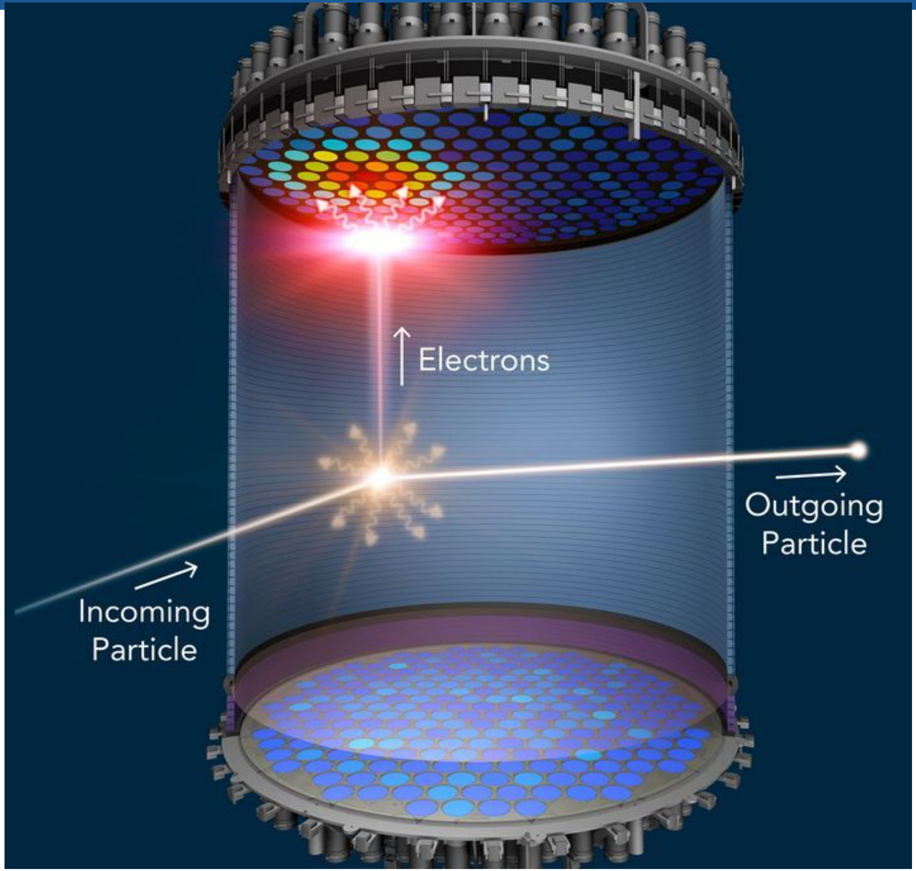


Image from Toomey (2020), *Smithsonian Magazine*; credit: SLAC National Accelerator Laboratory.

Dual-phase Time Projection Chamber (TPC)

Signals

S1: prompt scintillation in LXe

S2: proportional scintillation in GXe

Event reconstruction

X, Y: from PMT pattern

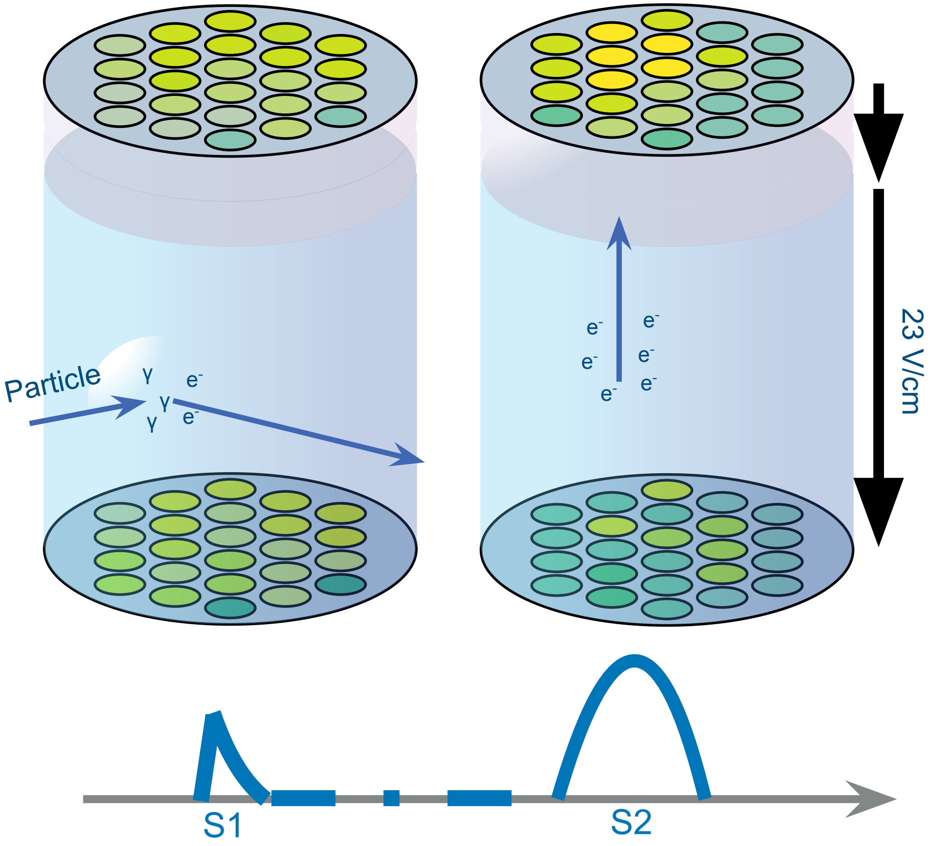
Z: drift time × drift velocity

Energy: from size of S1 and S2 signals

Particle discrimination using S2 to S1 ratio

Electronic recoil (ER): beta, gamma, …

Nuclear recoil (NR): neutron, WIMP



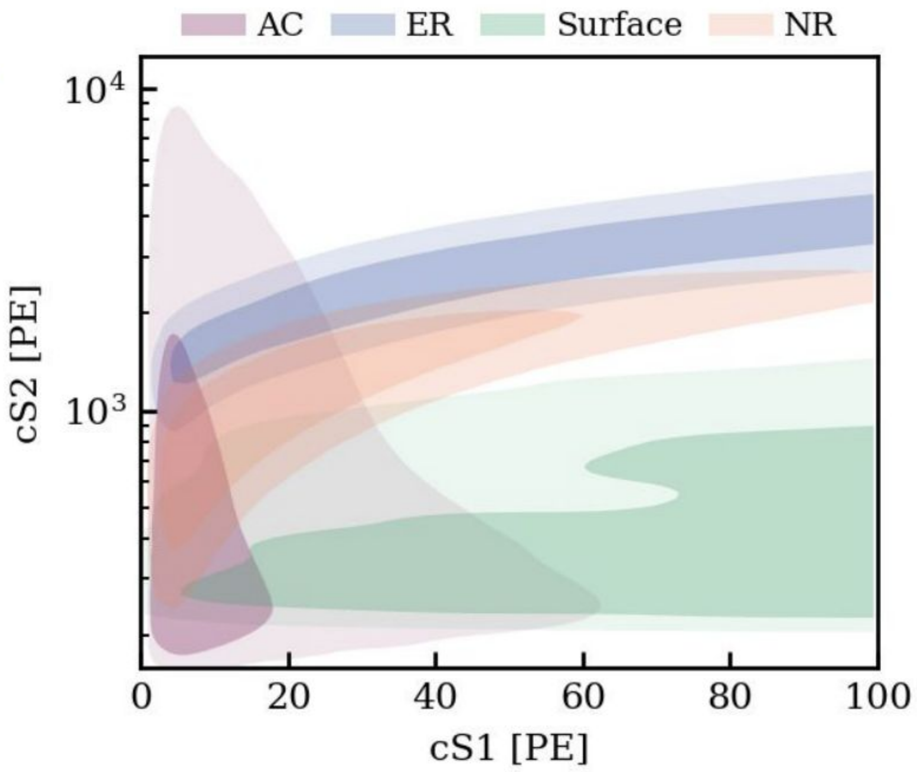
Source: Tan (2026, slide 22)

Differentiating Particles

- Neutron interactions look very similar to WIMPs
- **How do we account for this?**
  - Shielding
    - Location: XENON is under 1,400 meters of solid rock at the Laboratori Nazionali del Gran Sasso in Italy
    - Active Veto System: Neutrons cause Cherenkov radiation before entering detector
  - Differentiating
    - Neutrons scatter easily, WIMPs are unlikely to do so
    - S2/S1 ratio filters out electron(beta, gamma) vs nuclear(neutrons, WIMPs) recoils

Distribution

- A distribution of the detector background is crucial for identifying WIMPs
- This can be created using calibration data with simulation



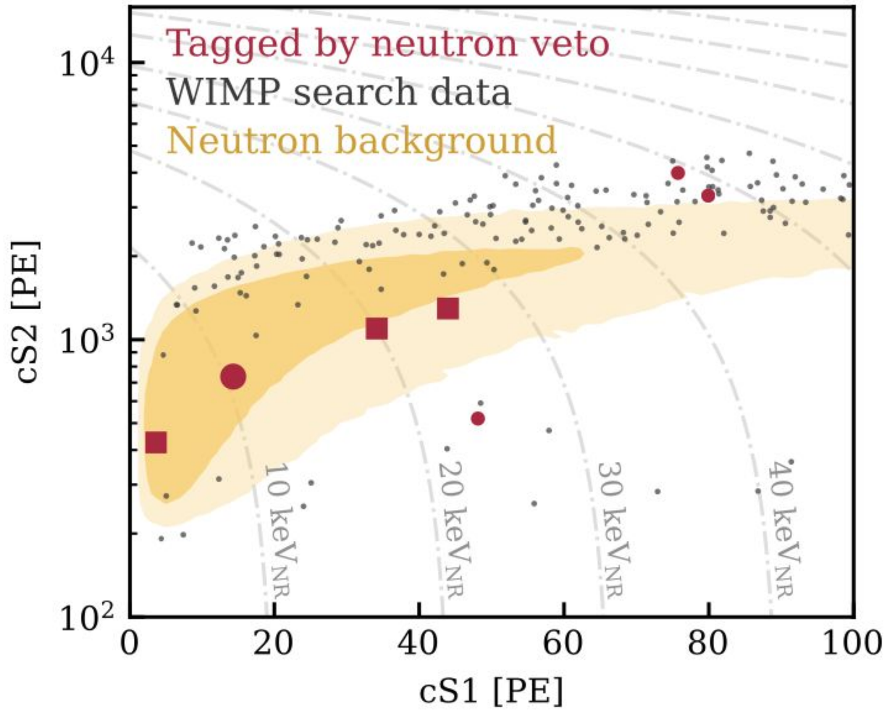
Source: Tan (2026, slide 26)

Simulation

- Full-chain neutron simulation uses Geant4 + FUSE
- Accurate, but too computationally expensive for repeated large-scale studies

Need for cheaper solution:

- Use full-simulation events to train a black box that predicts which neutron-scatter clusters contribute to the main reconstructed S2 peak



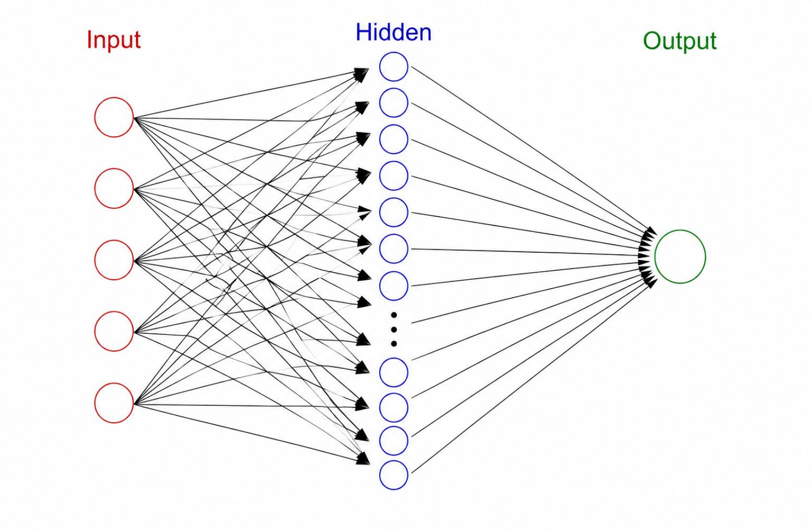
Source: Aprile et al. (2025, Fig. 7).

Data Structure

- **Event:** Neutron enters TPC
  - Interacts with xenon nucleus
- **Cluster:** Energy deposition releases cloud of electrons
  - There can be anywhere between 1-126 clusters in an event
  - Each cluster has 5 features
- **Target:** Calculate p_main for all clusters in an event

Neural Network Basics

- **Input:** 5 features per cluster
  - X
  - Y
  - Mean drift time
  - Drift Spread
  - Electron Count



- **Output:** P_Main per cluster
  - Percent contribution to the primary S2 peak

Problems with neural networks

- Events contain a variable number of clusters
- The topology of the entire event matters
  - Therefore, the model needs both cluster-level and event-level information.

- Ordering is generally not significant in particle physics

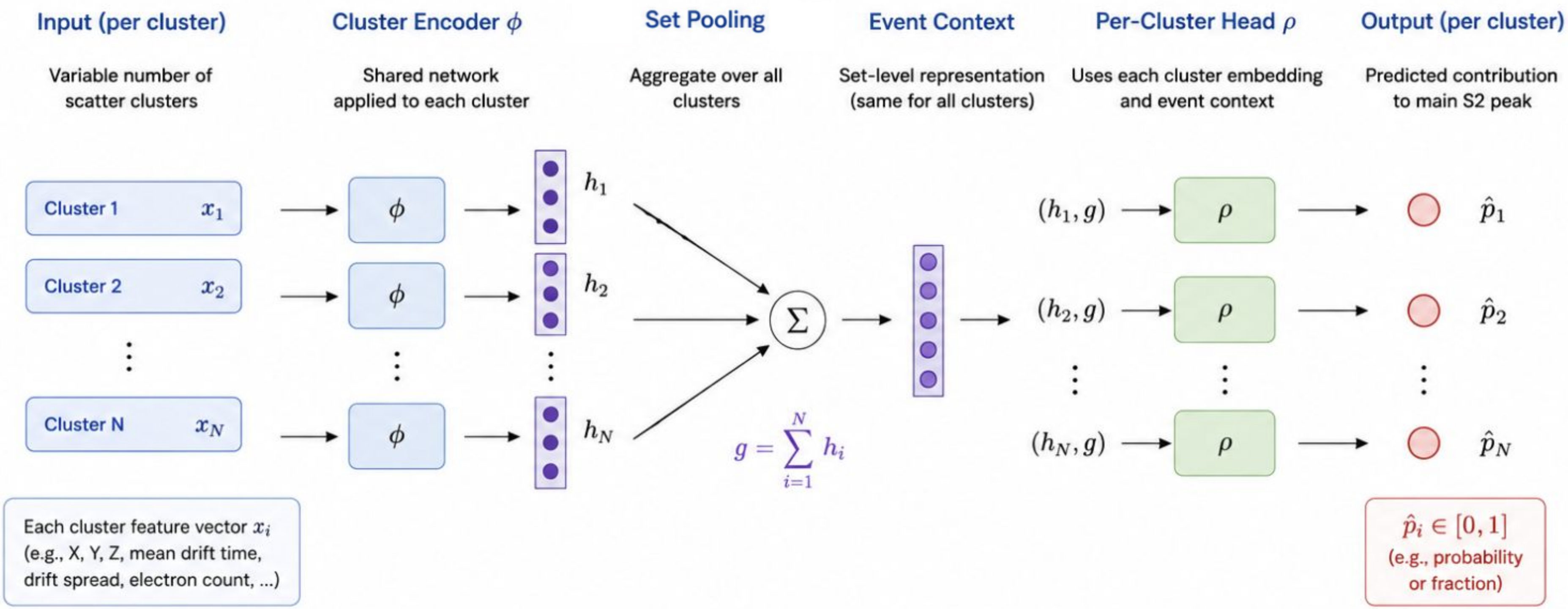
Introducing DeepSets

- Architecture for machine learning that addresses two crucial problems
  - Permutation Invariance
    - {A,B,C} = {B,C,A}
  - Set level context
    - Handles varying size data

$$f(\{x_i\}) = \rho\left(\sum_i \phi(x_i)\right)$$

(Zaheer et al., 2017)

Architecture

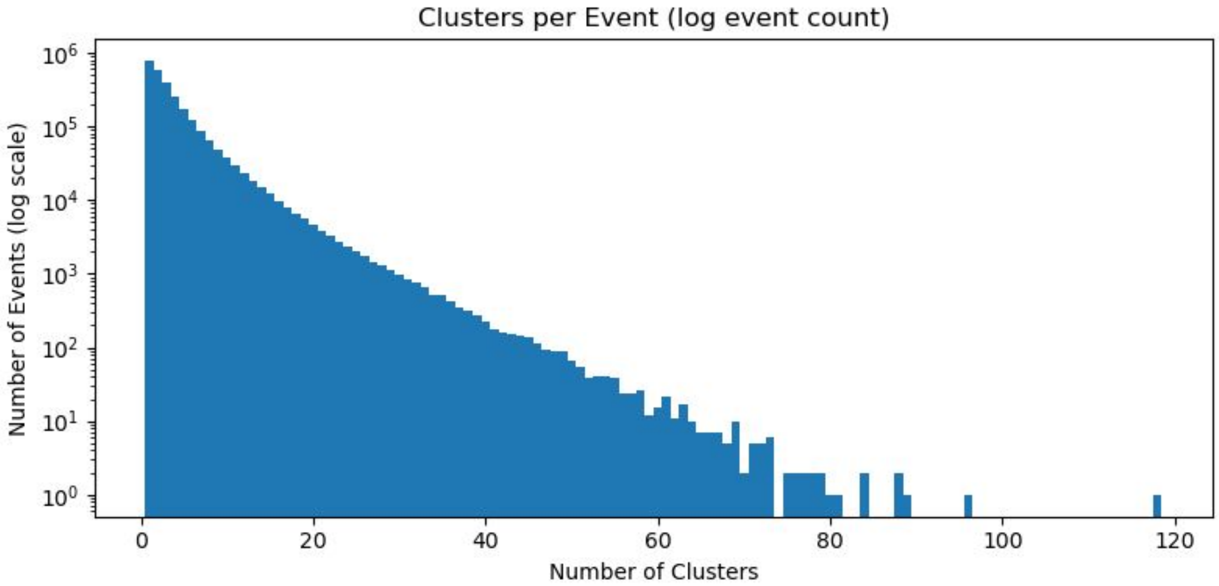


Training

- Neural network uses weights and biases to transform input features
- Test set output from neural net is evaluated against known values using a loss function
  - We use Mean Squared Error (MSE)
- Loss function is minimized using a gradient descent algorithm

Training Data

- 2.7 Million Events
  - 10.2 Million Clusters
- Skewed towards less clusters



Data Preprocessing

- Top 13 cm fiducial cut
  - Remove all events with clusters occurring in the top 13 cm of the TPC
  - ~660,000 events (~23%)
- Select train/test/validate split by events
  - Prevents contamination between sets
- Standardize features
  - Prevents unit conversions from dominating training

Baselines

Argmax

- Picks cluster with highest electron count and assigns it p_main = 1

Per Cluster MLP

- Simple machine learning algorithm that lacks set context

Metrics

Mean Squared Error

- average squared difference between a model's predicted values and the actual observed values.

Mean Absolute Error :

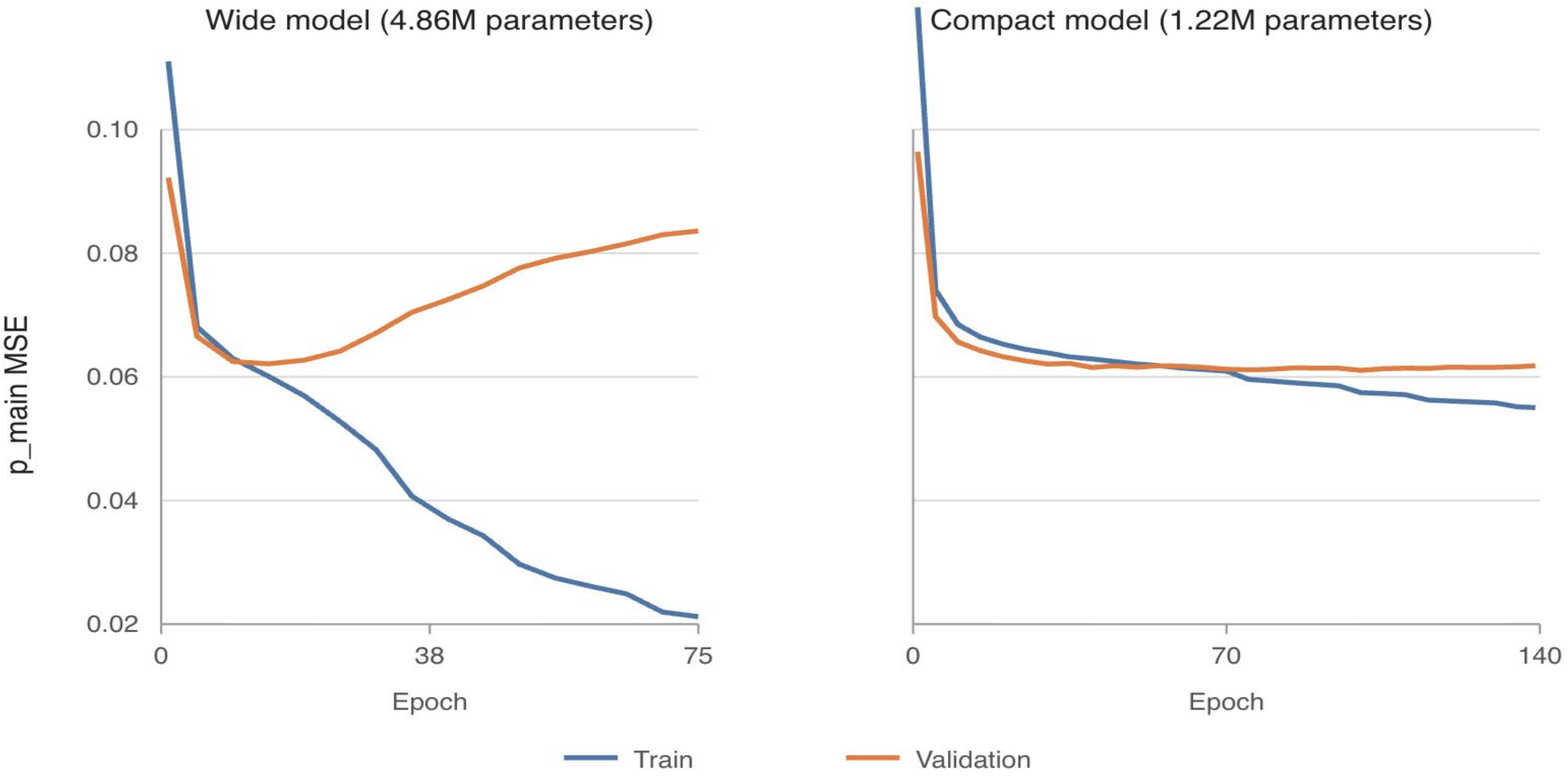
- the average size of errors between predicted values and actual values

Coefficient of Determination (R^2)

- the proportion of variation in a dependent variable that can be explained by an independent variable or model

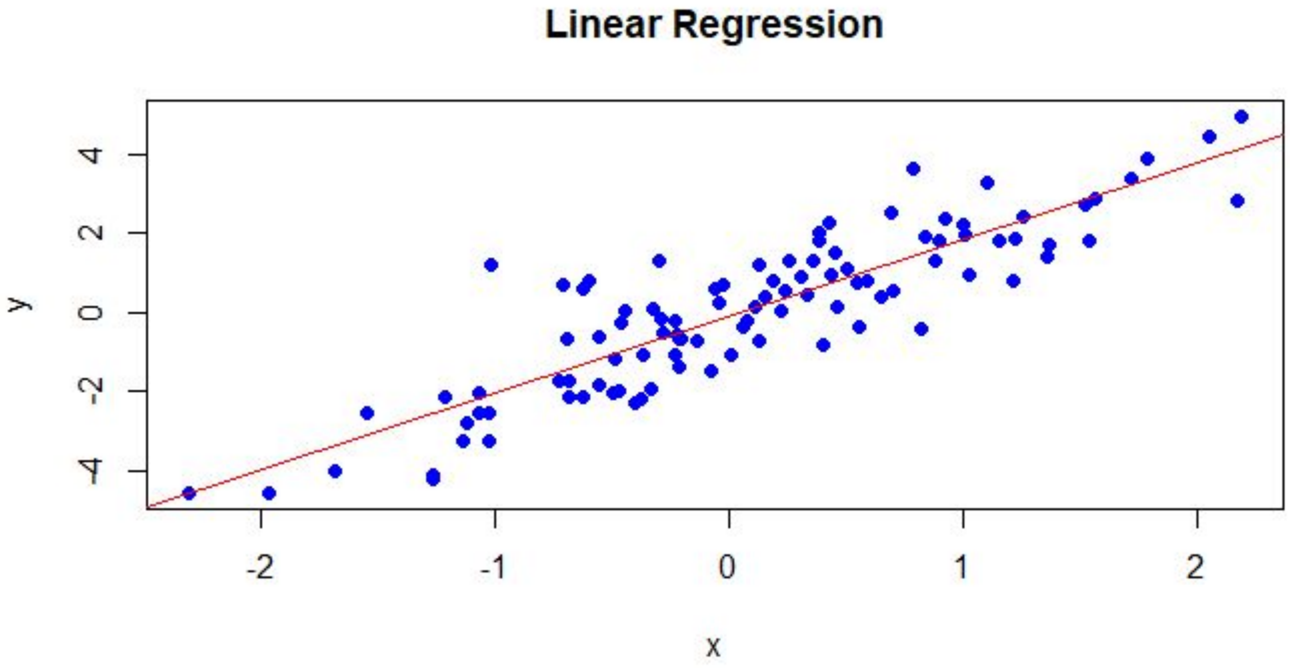
Tuning

- **Feature Engineering**
  - Adding 9 event level features derived from original 5
  - 6.85% reduction in MSE
- **Pooling**
  - Originally started with sum pooling → saw improvement with sum/mean/max pooling
  - Slight improvement with attention and gated sum pooling
- **Architecture sweeps**
  - Swept architecture configurations to find ideal width and depth

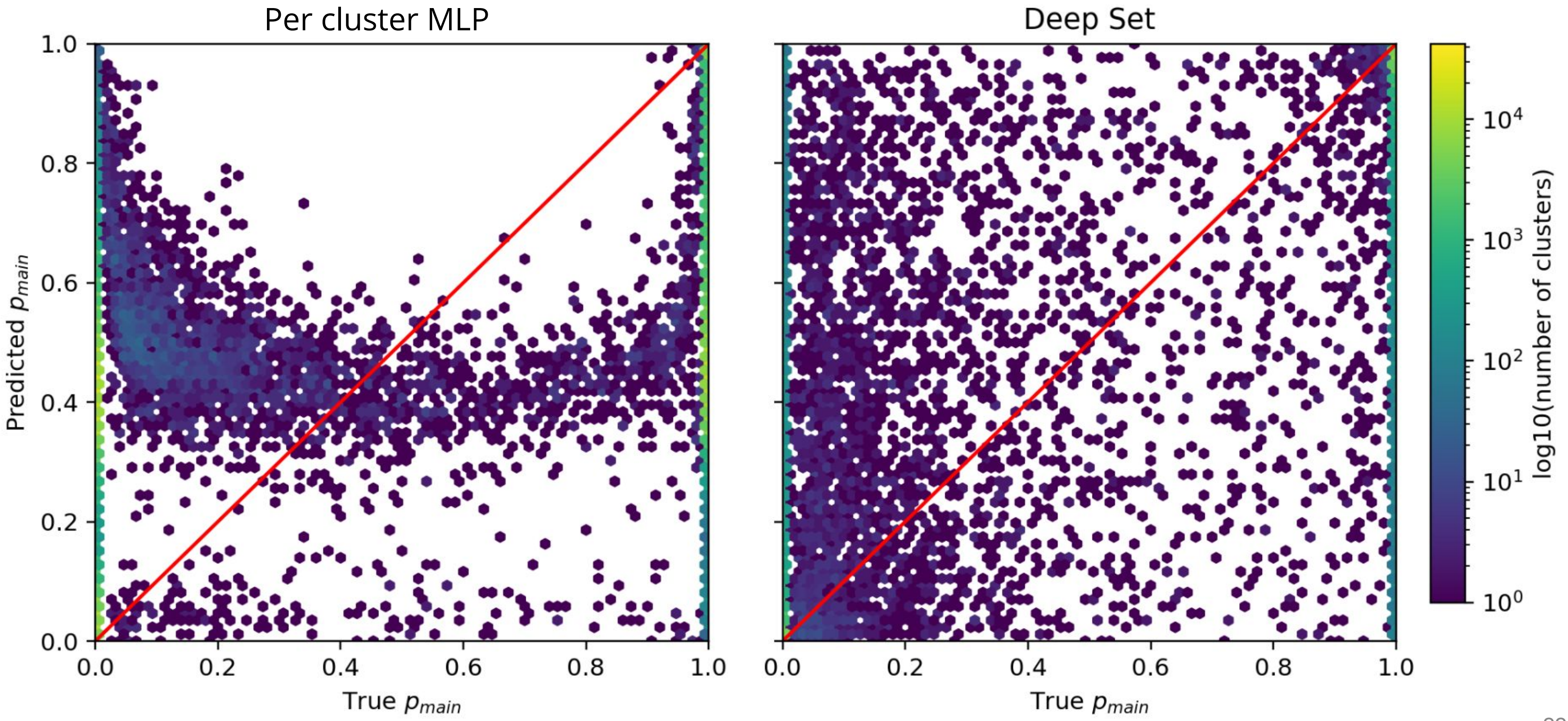

Wide model (4.86M parameters)
Compact model (1.22M parameters)
p_main MSE
0.10
0.08
0.06
0.04
0.02
0
38
75
0
70
140
Epoch
Epoch
Train
Validation

Model	Test MSE ↓	Test MAE ↓	Test R^2 ↑ [0,1]
Argmax	0.2293	0.2322	0.06
Per-cluster MLP	0.1862	0.3744	0.24
Deep Set	0.07	0.13	0.73
Deep Sets w/ Feature Engineering	**0.06**	**0.12**	**0.75**

What does a good result look like?



Results



Per cluster MLP
Deep Set
Predicted p_{main}
True p_{main}
True p_{main}
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
10^4
10^3
10^2
10^1
10^0
log10(number of clusters)

Conclusions

- Event level context and permutation invariance do provide meaningful gains to performance
- Model struggles with learning signals for a p_main value between 0 and 1 (2.5% of training data)
- Increase in model complexity will **not** yield better performance

Further exploration

- Audit failure cases - especially fractional p_mains
- Attempt training data tuning
  - Previous attempts at weighting
- Explore pooling options
- Simplify model and come at it with new knowledge

Acknowledgements

Thank you to Professor Elena Aprile, Dr. Pueh Leng Tan, the XENON collaboration, and Nevis Labs for their support

Citations

Aprile, E., et al. (2025). XENONnT WIMP search: Signal and background modeling and statistical inference. *Physical Review D*. Advance online publication. https://doi.org/10.1103//PhysRevD.111.103040

O'Hare, C. (2024, January 22). *Dark matter candidates* [Infographic]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Dark_matter_candidates.pdf

Tan, P. L. (2026, June 12). *Dark matter searches with liquid xenon* [PowerPoint slides]. REU Lectures Summer 2026, Columbia University.

Toomey, E. (2020, February 3). *New generation of dark matter experiments gear up to search for elusive particle*. *Smithsonian Magazine*.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Póczos, B., Salakhutdinov, R., & Smola, A. J. (2017). Deep sets. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 3391–3401).