

NuGraph3, the latest iteration of a Fermilab graph neural network, seeks to use simulated particle detector data to reconstruct neutrino interactions. To enhance the model, the clinic team focused on improving particle clustering, a major contributor to overall accuracy.

The model predicts particle labels by clustering detector hits (or nodes) that likely originated from the same particle. The clustering process assigns two embeddings to each node, which are used to pull similar nodes together, making the data easier to cluster.

In the original implementation, these embeddings were created immediately before starting the clustering process, after the model was done learning. Although they benefited from learned node features, the embeddings themselves were not incorporated into the model's core learning process. The team moved the initialization of the embeddings to the beginning of the model, then integrated the embeddings into the core, allowing them to be learned iteratively.

This change resulted in a significant improvement in the model's clustering accuracy, measured using a Rand score, as seen in the validation results below. However, the model training time increased as a trade-off. In the future, the team hopes to improve the efficiency of these improvements.

