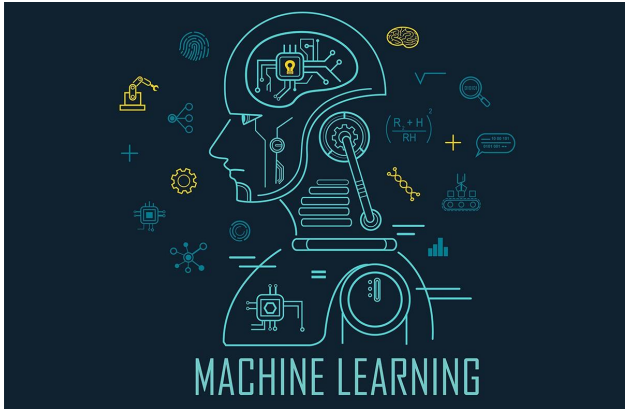
Optimization of Machine Learning Algorithms for LArTPCs

Katie Edwards
Iowa State University
Mentors: Dr. Jin Tian(ISU), Dr. Hongyang Gao(ISU), Dr. Jane Nachtman(UI)

Machine Learning Overview

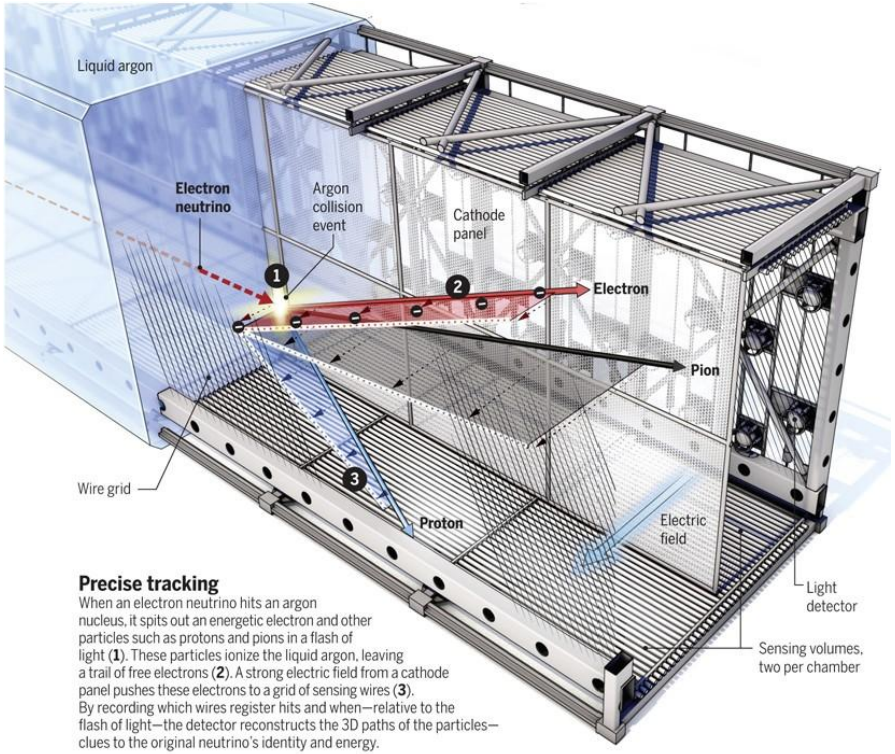


https://miro.medium.com/max/1400/1*c_fiB-YgbnMl6nntYGBMHQ.jpeg

- Used in many AI-like programs
- Completes tasks in a way that is similar to humans
- Can identify images or other patterns that they are trained to recognize
    - Good for analyzing neutrino interactions

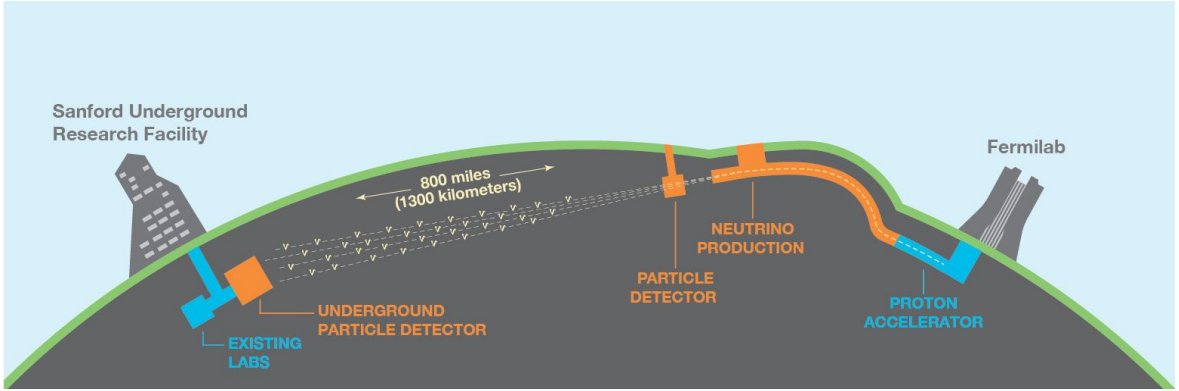
LArTPC

- Neutrinos have weak interactions which makes them hard to detect
- LArTPC is a particle imaging detector used to detect neutrino interactions
- Takes multiple 2D images using wires, or takes a 3D image using pixels
- Recent technology developed specifically for neutrino physics purposes



https://github.com/DeepLearnPhysics/lartpc_mlreco3d_tutorials/blob/master/book/Physics/lartpc.jpg

DUNE



https://www.dunescience.org/wp-content/uploads/2016/12/LBNE_Graphic_061615_2016.jpg

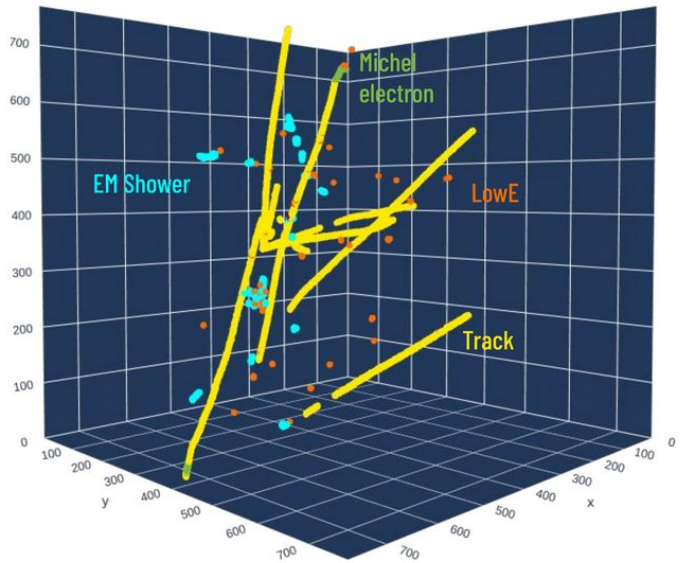
- Goal: to answer fundamental questions using neutrino analysis
  - Role of neutrinos and how they contribute to matter
  - Understand proton decay in nature to determine matter's stability
  - The role of neutrinos in black hole formation
- Two LArTPCs
  - Near Detector examines neutrinos before oscillation
  - Far Detector analyzes neutrinos after oscillation(detects flavor change)
  - Will use a total of 40,000 tons of liquid argon once built

Problem statement

- Algorithms vary in effectiveness
- Focusing on increasing accuracy in algorithms for better reconstruction results
- Sometimes it's more important to have accuracy in one field over another

```
Deghosting Accuracy: 0.9850
Segmentation Accuracy: 0.9954
PPN Accuracy: 0.8681
Clustering Accuracy: 0.9582
Clustering Edge Accuracy: 0.1975
Shower fragment clustering accuracy: 0.9733
Shower primary prediction accuracy: 1.0000
Track fragment clustering accuracy: 0.9976
Interaction grouping accuracy: 0.9869
Particle ID accuracy: 1.0000
Primary particle score accuracy: 1.0000
```

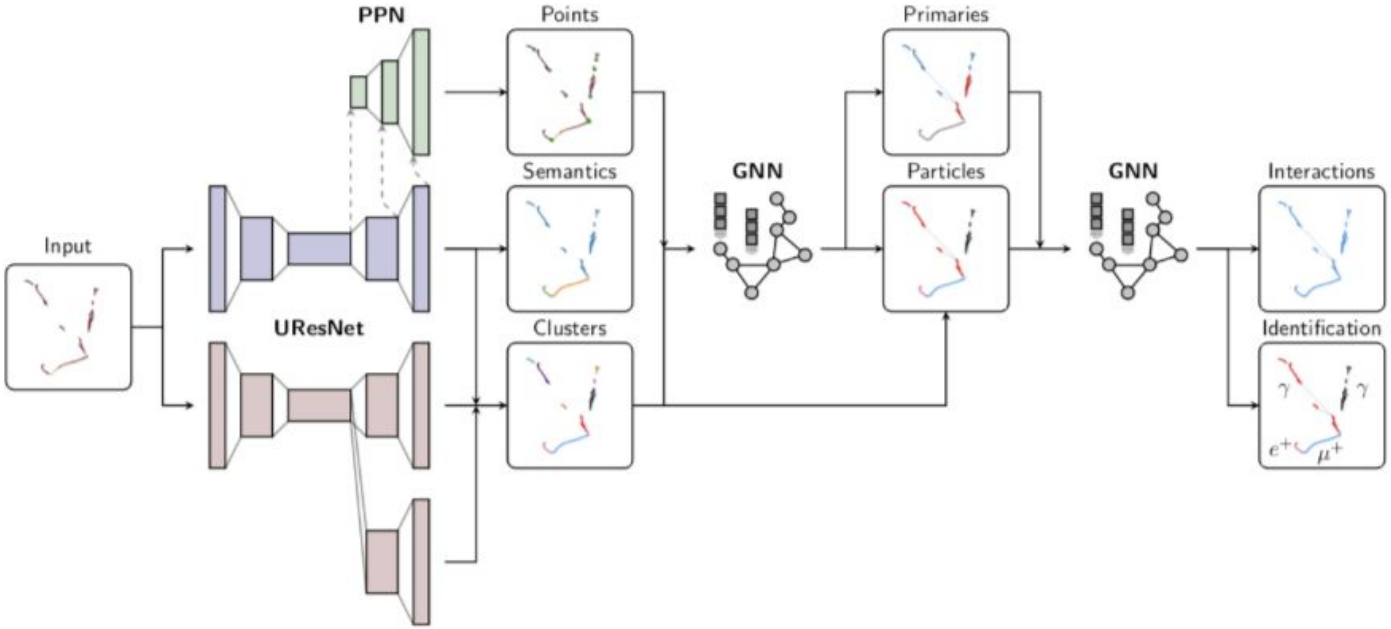
LArTPC MLReco3D



http://deeplearnphysics.org/lartpc_mlreco3d_tutorials/_images/semantic.png

- GitHub repository made for training machine learning models with images from LArTPCs
- Steps:
  - Semantic segmentation
  - Identifies points of interest
  - Clusters particles together
  - Identifies particles

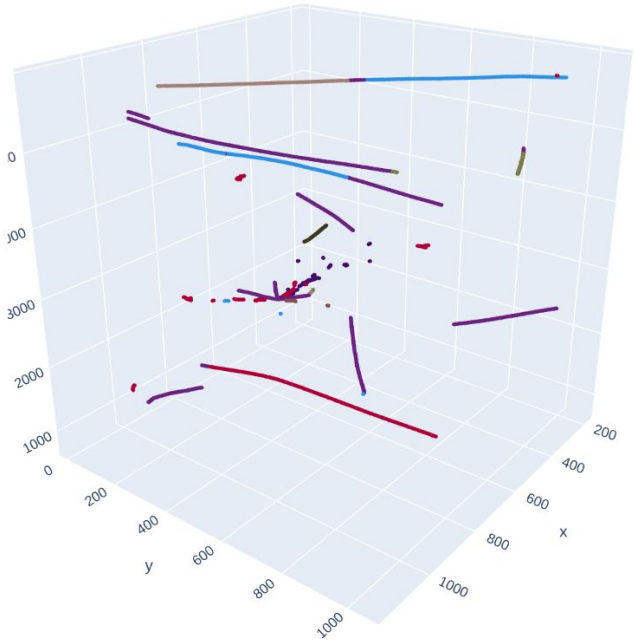
LArTPC MLReco3d Pipeline

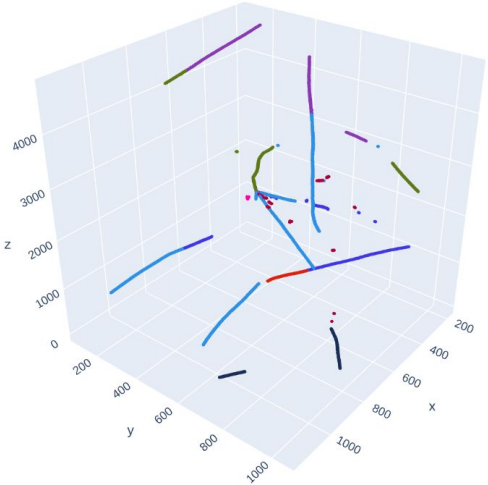


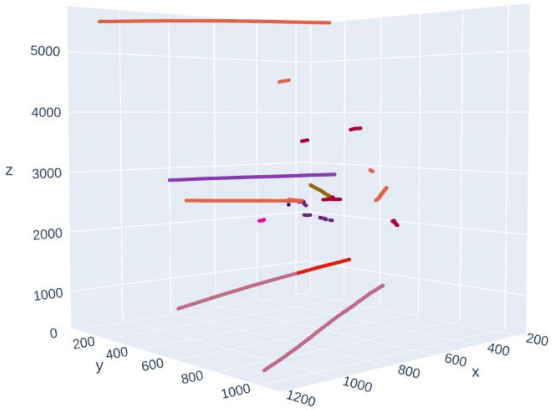
https://github.com/DeepLearnPhysics/larlite_mlreco3d/raw/develop/images/full_chain.png

Dataset

- Simulated data from the MINERvA experiment
- 10 different images used in training







Optimization Example

- Manipulated value that determines the radius that orphan particles get matched with neighboring clusters
- Original value was 1.0

jupyter cluster_graph_constructor.py 07/13/2022

Logout

File Edit View Language Python

```python
        # Initial Neighbor Graph Construction Mode
        mode = constructor_cfg.get('mode', 'knn')
        if mode == 'knn':
            self._init_graph = knn_graph
        elif mode == 'radius':
            self._init_graph = radius_graph
        else:
            raise ValueError('''Mode {} is not supported for initial
                graph construction!'''.format(mode))

        # Clustering Algorithm Parameters
        self.ths = constructor_cfg.get('edge_cut_threshold', 0.0) # Prob values 0-1
        if self.training:
            self.ths = 0.0
        # print("Edge Threshold Probability Score = ", self.ths)
        self.kwargs = constructor_cfg.get('cluster_kwargs', dict(k=5))
        # Radius within which orphans get assigned to neighbor cluster
        self._orphans_radius = constructor_cfg.get('orphans_radius', 1.0)
        self.use_cluster_labels = constructor_cfg.get('use_cluster_labels', True)

        # GraphBatch containing graphs per semantic class.
        if graph_batch is None:
            self._graph_batch = GraphBatch()
        else:
            assert graph_info is not None
            self._info = graph_info
            self._graph_batch = graph_batch
            self._num_total_nodes = self._graph_batch.x.shape[0]
            self._node_dim = self._graph_batch.x.shape[1]
            self._num_total_edges = self._graph_batch.edge_index.shape[1]

        self._edge_pred = None
        self._node_pred = None
```

Optimization Example

- Increases one field but decreases another field
  - Track fragment clustering accuracy clusters small track pieces together that are left by initial clustering algorithm
  - Interaction grouping accuracy predicts which edges should be kept
- The larger the radius value, the more orphan particles will get incorrectly matched to clusters that are far away, but more edge particles will be retrieved
- Sometimes some accuracy fields are more important than others

Before change:

```
Deghosting Accuracy: 0.9850
Segmentation Accuracy: 0.9954
PPN Accuracy: 0.8681
Clustering Accuracy: 0.9582
Clustering Edge Accuracy: 0.1975
Shower fragment clustering accuracy: 0.9733
Shower primary prediction accuracy: 1.0000
Track fragment clustering accuracy: 0.9976
Interaction grouping accuracy: 0.9869
Particle ID accuracy: 1.0000
Primary particle score accuracy: 1.0000
```

Radius = 10.0:

```
Deghosting Accuracy: 0.9850
Segmentation Accuracy: 0.9954
PPN Accuracy: 0.8681
Clustering Accuracy: 0.9582
Clustering Edge Accuracy: 0.1975
Shower fragment clustering accuracy: 0.9733
Shower primary prediction accuracy: 1.0000
Track fragment clustering accuracy: 0.9905
Interaction grouping accuracy: 0.9882
Particle ID accuracy: 1.0000
Primary particle score accuracy: 1.0000
```

Optimization Example

- Causes a similar increase in interaction grouping accuracy, but decreases a different accuracy field than before
  - Primary particle score accuracy classifies which particles are closest to the vertex
- When the radius decreases, more primary particles might be misidentified as non-primary

Before change:

```
Deghosting Accuracy: 0.9850
Segmentation Accuracy: 0.9954
PPN Accuracy: 0.8681
Clustering Accuracy: 0.9582
Clustering Edge Accuracy: 0.1975
Shower fragment clustering accuracy: 0.9733
Shower primary prediction accuracy: 1.0000
Track fragment clustering accuracy: 0.9976
Interaction grouping accuracy: 0.9869
Particle ID accuracy: 1.0000
Primary particle score accuracy: 1.0000
```

Radius = 0.5:

```
Deghosting Accuracy: 0.9850
Segmentation Accuracy: 0.9954
PPN Accuracy: 0.8681
Clustering Accuracy: 0.9582
Clustering Edge Accuracy: 0.1975
Shower fragment clustering accuracy: 0.9733
Shower primary prediction accuracy: 1.0000
Track fragment clustering accuracy: 0.9976
Interaction grouping accuracy: 0.9879
Particle ID accuracy: 1.0000
Primary particle score accuracy: 0.9688
```

Conclusion

- Machine learning is very useful for analyzing neutrino interactions
- The better the models are, the more accurate predictions will be
- The more machine learning models are trained now, the better the reconstruction models will perform, which will lead to more ground-breaking discovery in particle detection

Thank you!

Thank you to IRIS HEP, my three mentors: Jane Nachtman(U of I), Hongyang Gao(ISU), and Jin Tian(ISU), and special thanks to the University of Iowa DUNE group: Orgho Neogi, Gurkan Karaman, Jacob Andrews, and Avi Kaufman

References

DeepLearnPhysics. "DeepLearnPhysics/lartpc_mlreco3d." *GitHub*, https://github.com/DeepLearnPhysics/lartpc_mlreco3d.

Fermilab. (n.d.). *DUNE*. Deep Underground Neutrino Experiment. https://www.dunescience.org/.

Getting Started — Lartpc_mlreco3d Documentation. https://lartpc-mlreco3d.readthedocs.io/en/latest/GettingStarted.html. Accessed 16 July 2022.

Getting Started — Lartpc_mlreco3d Tutorials. http://deeplearnphysics.org/lartpc_mlreco3d_tutorials/index.html. Accessed 16 July 2022.

Ju, Xiangyang, et al. "Performance of a Geometric Deep Learning Pipeline for HL-LHC Particle Tracking." *The European Physical Journal C*, vol. 81, no. 10, Oct. 2021, p. 876. *Springer Link*, https://doi.org/10.1140/epjc/s10052-021-09675-8.

"Machine Learning, Explained." *MIT Sloan*, https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained. Accessed 27 July 2022.

Psihas, Fernanda, et al. "A Review on Machine Learning for Neutrino Experiments." *International Journal of Modern Physics A*, vol. 35,no. 33, 2020, p. 2043005., https://doi.org/10.1142/s0217751x20430058.

The Dune Collaboration. *Separation of Track– and Shower–like Energy Deposits in ProtoDUNE-SP Using a Convolutional Neural Network.* 31 Mar. 2022.