Yo what's up my Alpha-G

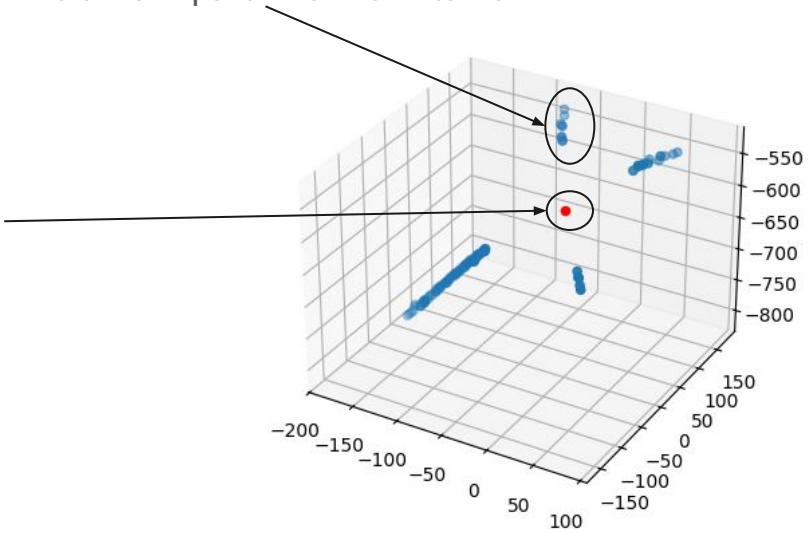
Atakan A. (SFU/TRIUMF)
Nikita D. (SFU/TRIUMF)
Surya R. (UBC/TRIUMF)

*Aah yes humour, we have it too

The experiment

“Does antimatter respond differently to gravity?”

- The idea : trap antihydrogen and let it fall under gravity.
- As antihydrogen falls, it annihilates upon contact with matter
  - Vestiges of its existence are found in tracks of particles left in its wake



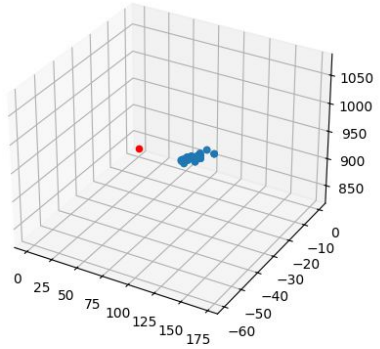
Vertex Reconstruction

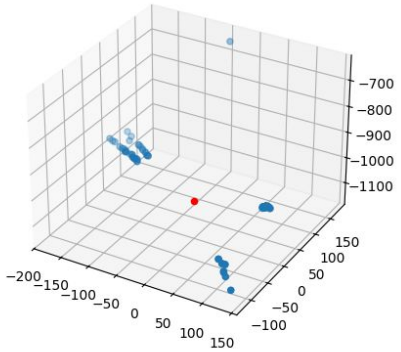
If we can trace back from the tracks and figure out where antihydrogen annihilated, we’ll be able to quantify if antihydrogen “fell” differently compared to ordinary boring vanilla hydrogen

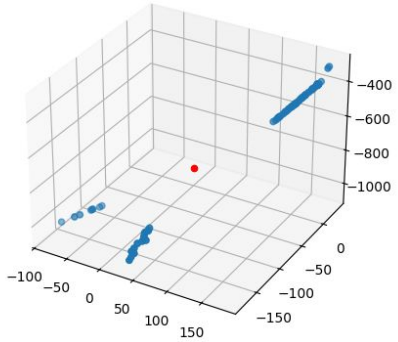
Data Exploration

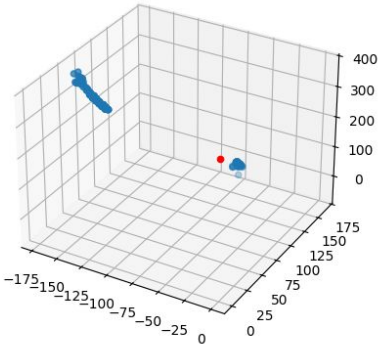
Points of consideration :

- Is the data skewed?
  - Models are essentially code, so they are inherently dumb.
  - We don't want the model to learn about just a subset of possible vertex locations.
- Are the data points of the same size?
  - Each event in the dataset is just a collection of 3D points (the vertex and the "point cloud"/tracks)
  - Do all the point clouds have the same size?







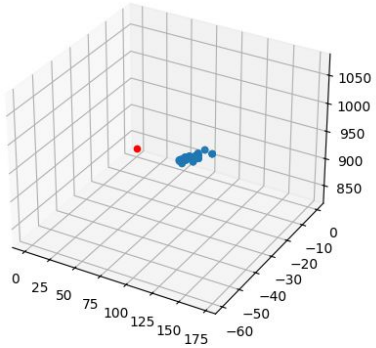


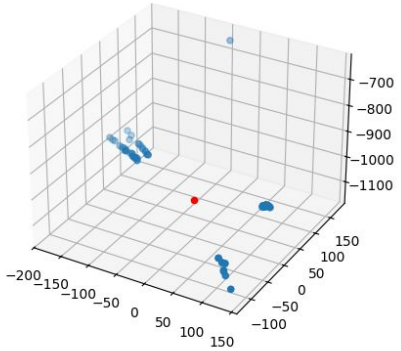
Data Exploration

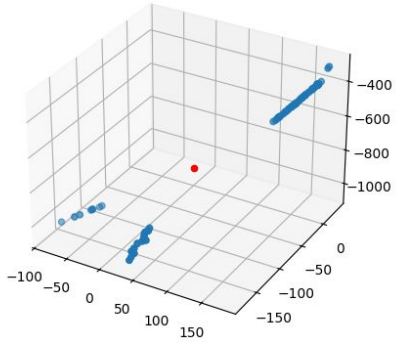
Points of consideration :

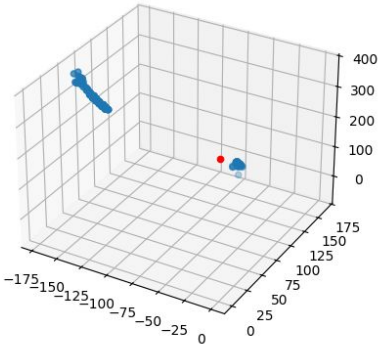
-

- Are the data points of the same size?
  - Each event in the dataset is just a collection of 3D points (the vertex and the “point cloud”/tracks)
  - Do all the point clouds have the same size?









Data Exploration

Points of consideration :

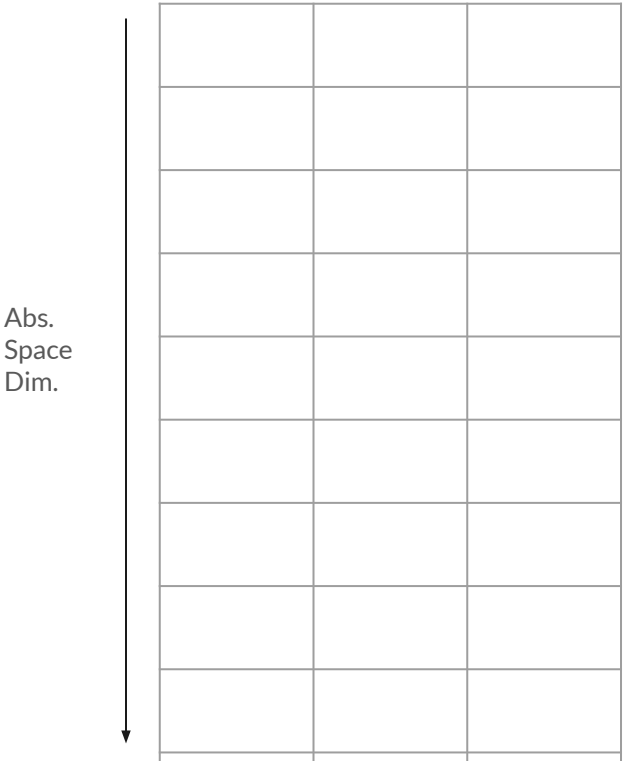
- Is the data skewed?
  - Models are essentially code, so they are inherently dumb.
  - We don't want the model to learn about just a subset of possible vertex locations.

Data Sorting

Vertex : Select out the z component for training (this is the eventual prediction)

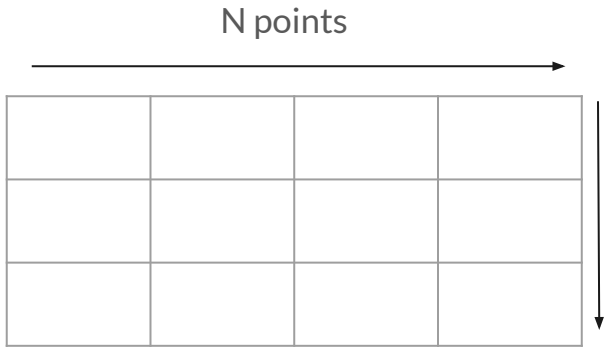
PC : Sort each point based on x-y distance from origin, and set a fixed length of no. of points allowed for each event (if event has lesser points than this max length, just wrap the array around itself)

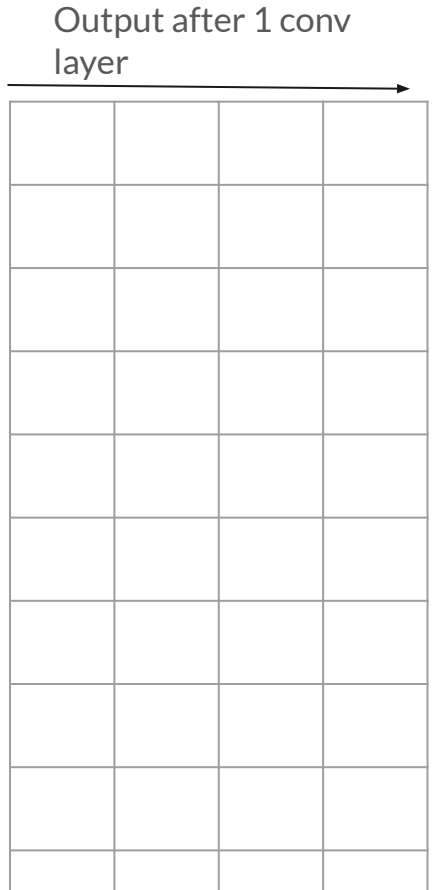
The Inner Model



Feature Extraction

Each point (with 3 dimensions) in a given event is mapped to an abstract higher dimensional space (64 for eg.) using a series of 1D conv layers (kernel size ; with batch normalisation and and ReLU activation)





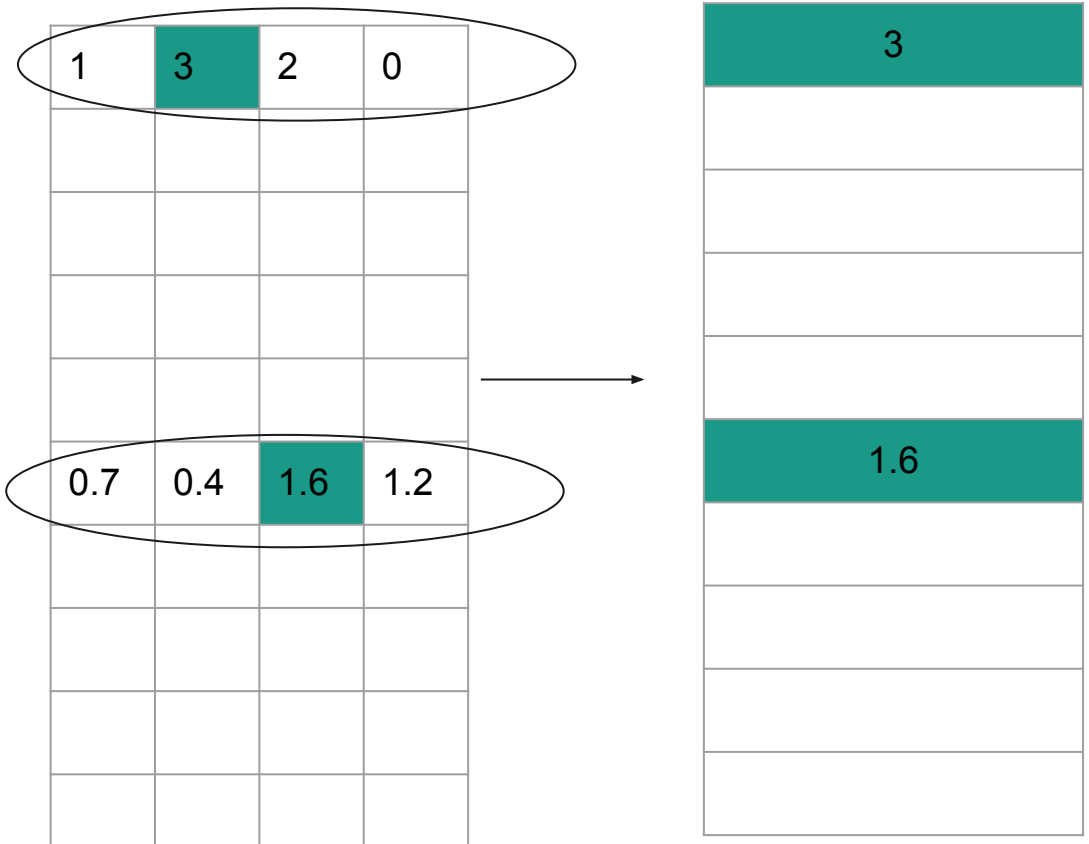
The Inner Model (Contd.)

Max Pooling

- Need to compress output from all conv layers (which is a ginormous matrix essentially) into fixed size array with most important features.
- Permutation invariant

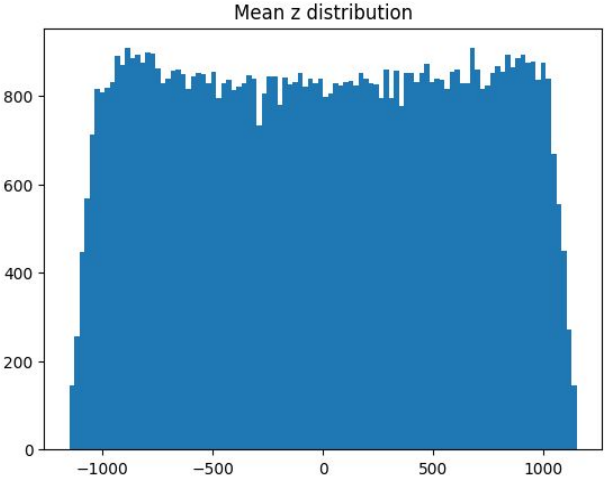
Regression

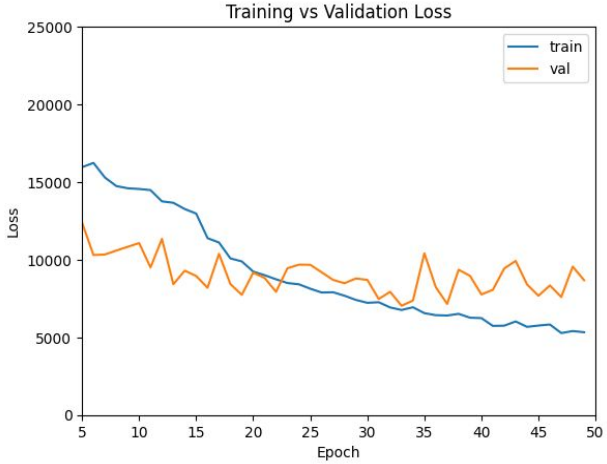
Fully connected linear layers that map this max pooled output to a single number (prediction for z coordinate of vertex)



Baseline Training Results

- Dataset was split into train, val and test subsets by shuffling and extracting data
- K-fold validation might be done in future



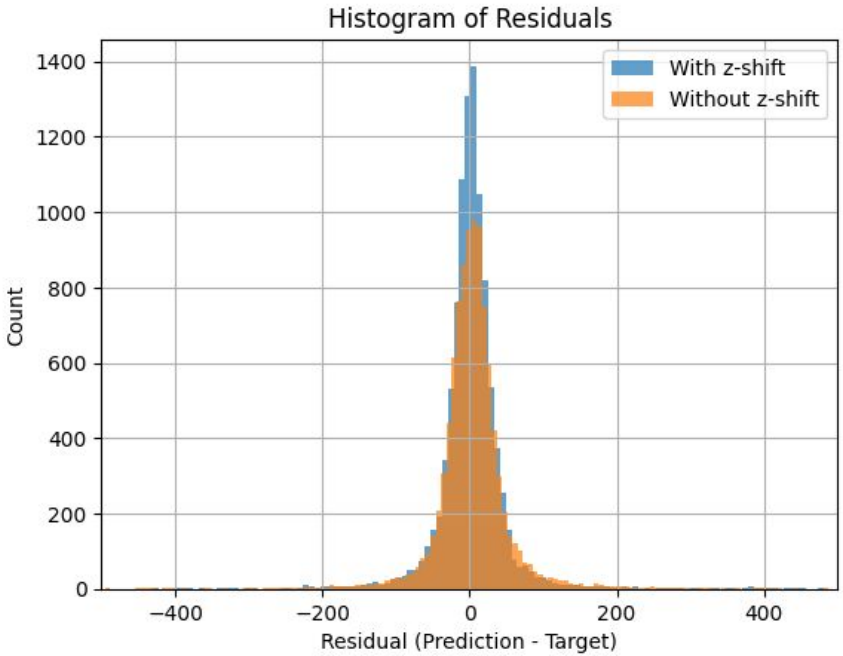


Z Shift : aka allocate the boring tasks to a grad student

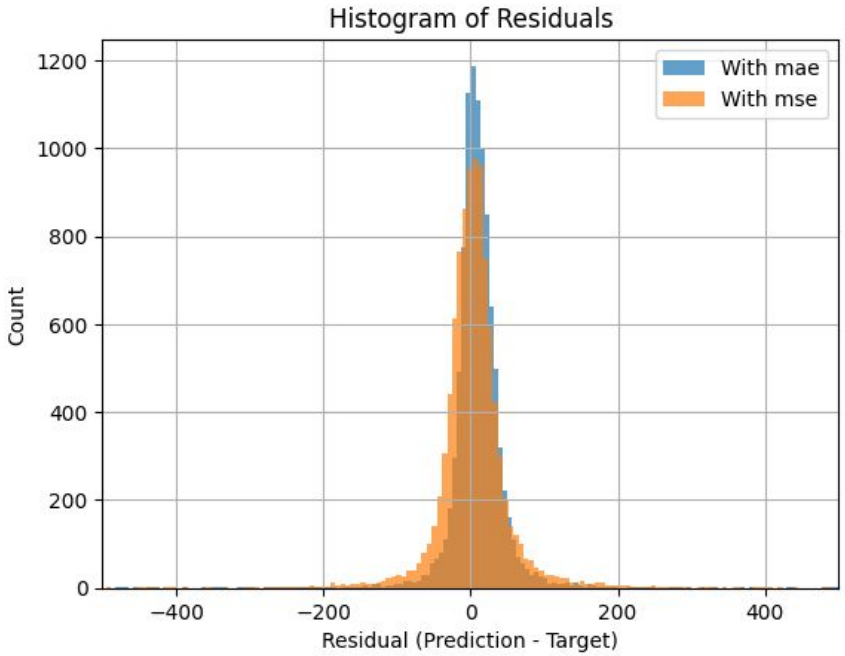
- Idea : one network **learning to predict a z value** is humbled by two networks trying to learn fine-tuned details about the z value
- The first internal network makes a first naive guess about where the predicted z value is
- This **“z shift” is subtracted from every point** in every event in the dataset, including the vertex
- A second network uses this “shifted” dataset to come up with a further fine tuned prediction for where the actual vertex should be
- More the merrier

Z-Shift Modification and different loss functions

Mean Squared Error



Mean Absolute Error



Things that one could do

- Performing a proper hyperparameter scan
- Mess around with loss function and optimiser
- Use a much larger dataset
- **A different network altogether? (aka the hubris of youth)**

Alternative Architecture: The DeepSets[1]

- Main ideas:

1) Encode with ϕ(x) network: Affine transform [x, y, z] → [—1024 dim —]
2) Aggregate ∑ , (sum/mean/max): We used **max pooling** over the points
   - Data shape (Batch, NumPoints, EmbedDim) → (Batch, EmbedDim)
3) Final ρ(x) network for final decision: (Batch, EmbedDim) → (Batch,)

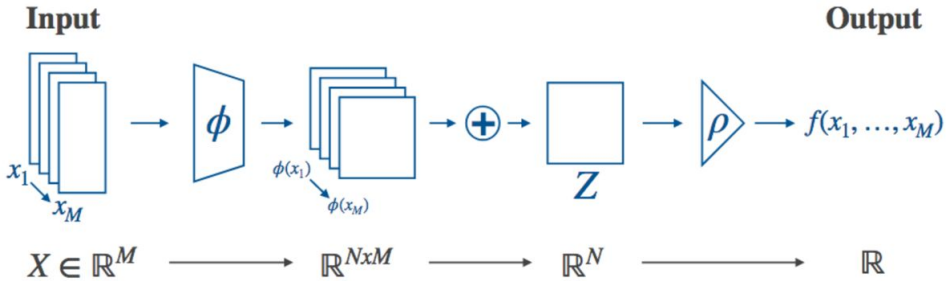
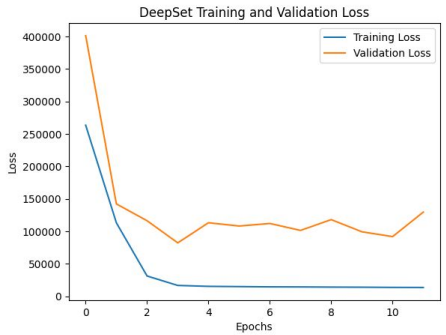
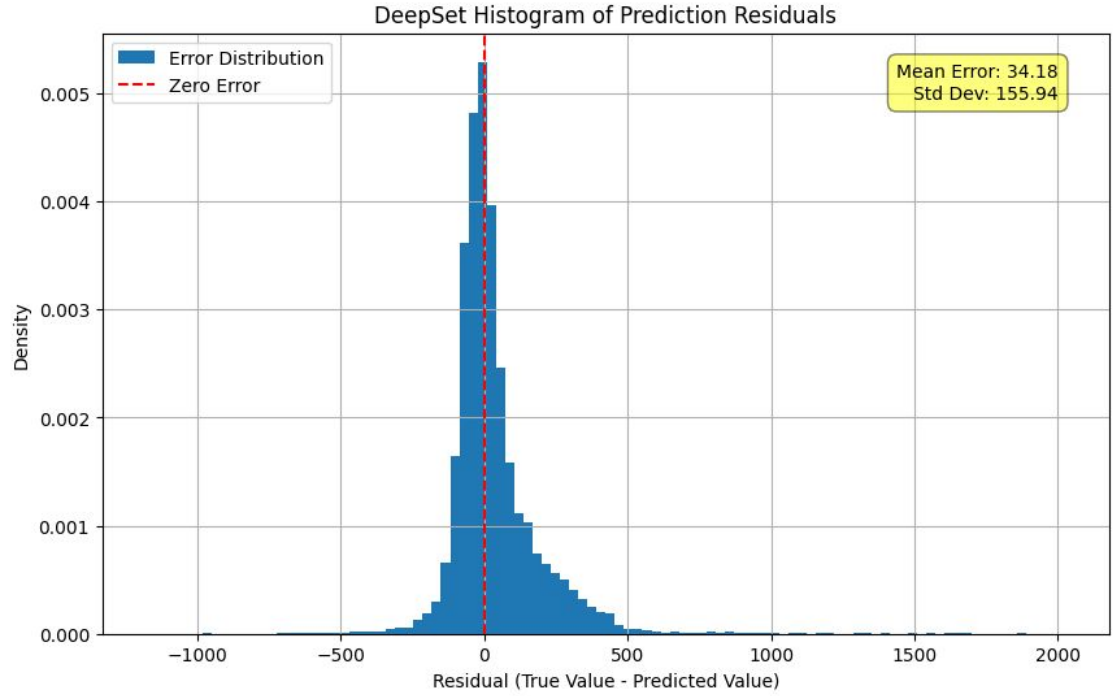


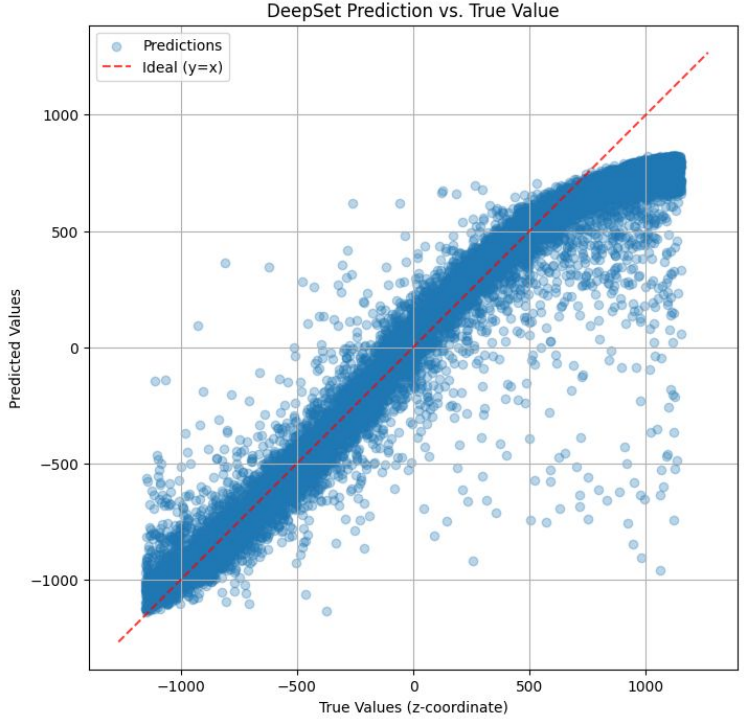
Figure Source

[1]https://arxiv.org/pdf/1703.06114

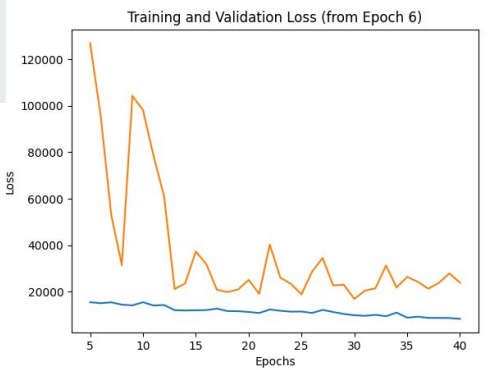
DeepSet Results

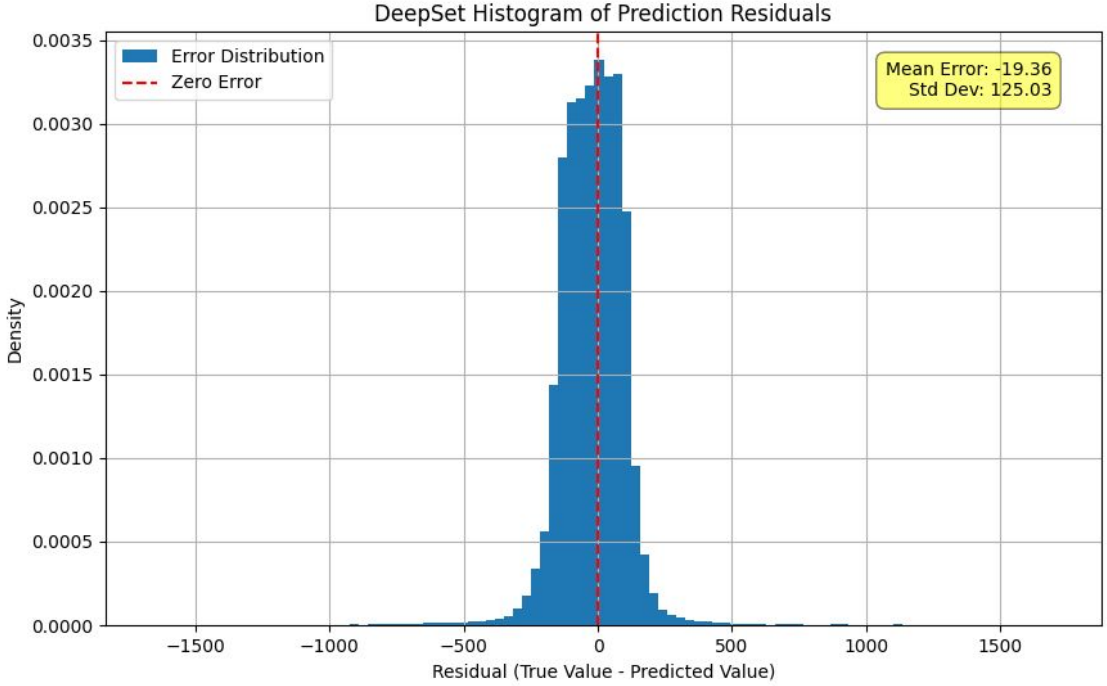


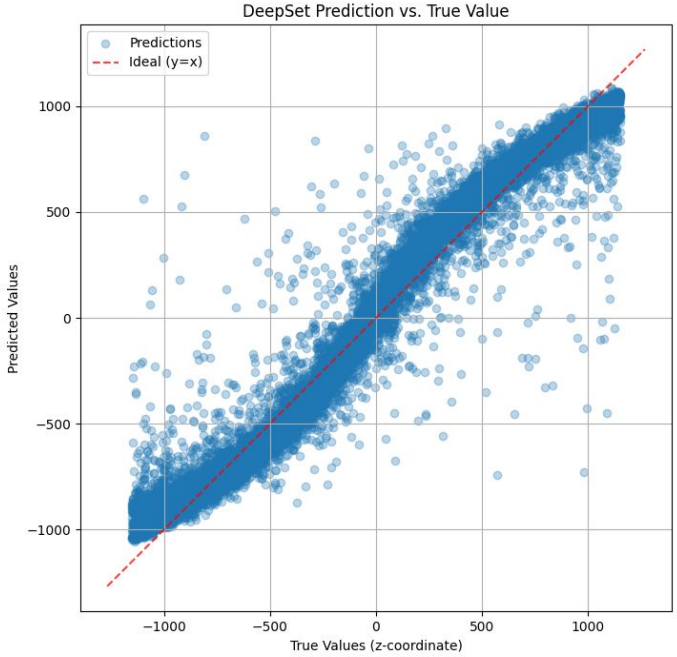




Gigantic DeepSet With Residual Connections







PEACE