Reconstructing HIBEAM Event Vertex using Machine Learning Techniques

Yamna Shaikh

April 15, 2026

One of the most important open questions in modern physics is the observed imbalance between matter and antimatter in the Universe. According to current theoretical understanding, the Big Bang should have produced matter and antimatter in equal amounts. However, the observable Universe is strongly dominated by matter, indicating the presence of physical processes beyond the Standard Model. A key ingredient in explaining this asymmetry is the violation of baryon number conservation. One of the possible mechanisms to observe such a violation is neutron-antineutron oscillations ($n \to \bar{n}$), a process in which a neutron spontaneously converts into an antineutron.
The HIBEAM (High-Intensity Baryon Extraction and Measurement) program at the European Spallation Source (ESS) is a proposed experiment designed to search for neutron-antineutron oscillations with high sensitivity. The signal is expected to appear as a rare annihilation event measured by the Time Projection Chamber (TPC) of HIBEAM. The measured particle trajectories are used to reconstruct the interaction vertex.
The primary goal of this thesis is to reconstruct the event vertex for the proposed HIBEAM experiment using simulated detector data. Each event is given as a set of detector hits with spatial and time coordinates (x, y, z, t). The data set includes smearing for detector resolution and background contributions from Compton electrons. We estimate the interaction vertex (x_v, y_v) using two reconstruction approaches. First, a classical geometry-based pipeline is implemented, where the hits are connected using a k-Nearest-Neighbor (kNN) graph, grouped into candidate tracks using connected-component clustering. Finally, the track directions are approximated using Principal Component Analysis (PCA) and extrapolated to the target plane to obtain a vertex estimate. Second, a graph-based machine learning approach is implemented where, each event is represented as a graph of hits and the DynEdge Model of Graph Neural Network(GNN) is trained to obtain the vertex coordinates.
Both methods are evaluated using residuals between predicted and truth vertex coordinates and the fraction of successfully reconstructed events. In conclusion, the GNN approach provides more stable and accurate vertex estimates, while the kNN and PCA pipeline is more sensitive to background contamination, resulting in larger outliers and a wider residual distribution.