**LiDAR Wire Clustering: Algorithmic Approach**

The goal of this project is to identify individual powerline wires from 3D LiDAR point clouds. The datasets provided present unique challenges, including varying point densities, large gaps, and noise. Through an iterative process, we determined that no single algorithm could solve all cases. The final solution is a **multi-strategy approach** that applies a specific, fine-tuned algorithm for each type of dataset.

This document outlines the journey from a simple approach to the final, robust solution.

**Attempt 1: Global Clustering**

The most straightforward approach is to apply a powerful, general-purpose clustering algorithm to the entire point cloud at once.

* **Strategy**: Use a density-based algorithm like **DBSCAN** or **HDBSCAN** on all points. These algorithms are excellent at finding arbitrarily shaped clusters and are robust to noise.
* **Result**: This worked for the cleanest (Easy) dataset but failed on all others. For datasets with large gaps or varying densities, it either grouped everything into one cluster or classified most of the wires as noise.
* **Conclusion**: A "one-size-fits-all" algorithm is not sufficient. The specific problems of each dataset must be addressed individually.

**Attempt 2: The Segment-and-Trace Method**

Inspired by academic research, we developed a more sophisticated "divide and conquer" method.

* **Strategy**:
  1. Slice the point cloud into thin vertical segments along its primary axis.
  2. In each slice, run a local DBSCAN to find the circular cross-sections of the wires.
  3. Trace each wire from one slice to the next by connecting the closest centroids.
* **Result**: This was a major breakthrough. It worked perfectly on the Easy dataset, as it could follow the continuous flow of points. However, it struggled on other datasets:
  1. On datasets with many close wires (Medium), the trace would sometimes "jump" to an adjacent wire.
  2. On datasets with large gaps (Hard), the trace would terminate because it couldn't find a close-enough point to connect to.
* **Conclusion**: This method is powerful but sensitive. Its success depends heavily on the quality and continuity of the data.

**Attempt 3: The Pre-Alignment Breakthrough**

We observed that the Segment-and-Trace method assumed the wires were parallel to the X-axis. This was not always the case.

* **Strategy**: Add a critical **preprocessing step**. Before any clustering, use **Principal Component Analysis (PCA)** to calculate the dominant direction of the entire point cloud and rotate it so the wires are perfectly aligned with the X-axis.
* **Result**: This dramatically improved the results for the Medium and Extra-Hard datasets. After alignment, the simple **Global DBSCAN** (Attempt 1) was now able to perfectly separate the distinct, parallel wires. However, this alignment made the result for the Easy dataset *worse*, as the now-perfectly-regular spacing confused the Segment-and-Trace logic.
* **Conclusion**: Pre-alignment is a powerful and necessary tool, but it's not universally beneficial. This confirmed that a tailored approach is required.

**The Final Multi-Strategy Solution**

Our final, successful pipeline combines all these learnings. A master function analyzes each dataset and directs it to one of three specialized strategies:

1. **For the Easy Dataset**: The original **Segment-and-Trace** method is used on the **unaligned** data. This remains the most effective strategy for this clean, simple case.
2. **For the Medium and Extra-Hard Datasets**: These datasets benefit most from order and separation.
   * First, the data is **pre-aligned** using PCA.
   * Then, a **Global DBSCAN** is applied. With the data aligned, this simple density-based approach is now powerful enough to separate the wires cleanly, even with the noise present in the Extra-Hard file.
3. **For the Hard Dataset**: This dataset's primary challenge is extreme gaps, which break both global density and simple tracing.
   * First, the data is **pre-aligned** using PCA.
   * Then, a modified **Segment-and-Trace** strategy is used with **extremely aggressive gap-jumping parameters**. This allows the trace to "leap" across large empty spaces to connect the sparse wire segments, successfully reconstructing the wires where other methods fail.