Main Cluster properties:

1. Homogenous within cluster.
2. Heterogenous among the clusters.

Task for given data:

1. Data Cleaning
2. Stats Correlation
3. Preprocessing
4. Normalization
5. Use Algorithm
6. Compare before and after

In diabetes dataset, use DT, KNN and Linear Regression

Research Paper

1. “BERT”
2. “Attention is all you need”