



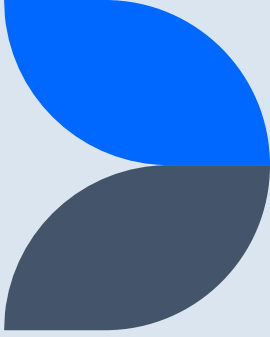
HDBSCAN

Clustering Algorithm

Machine Learning



Agenda



01

Introduction

02

HDBSCAN

03

What is
HDBSCAN

04

Key
Concepts:

05

How
HDBSCAN
Works:

06

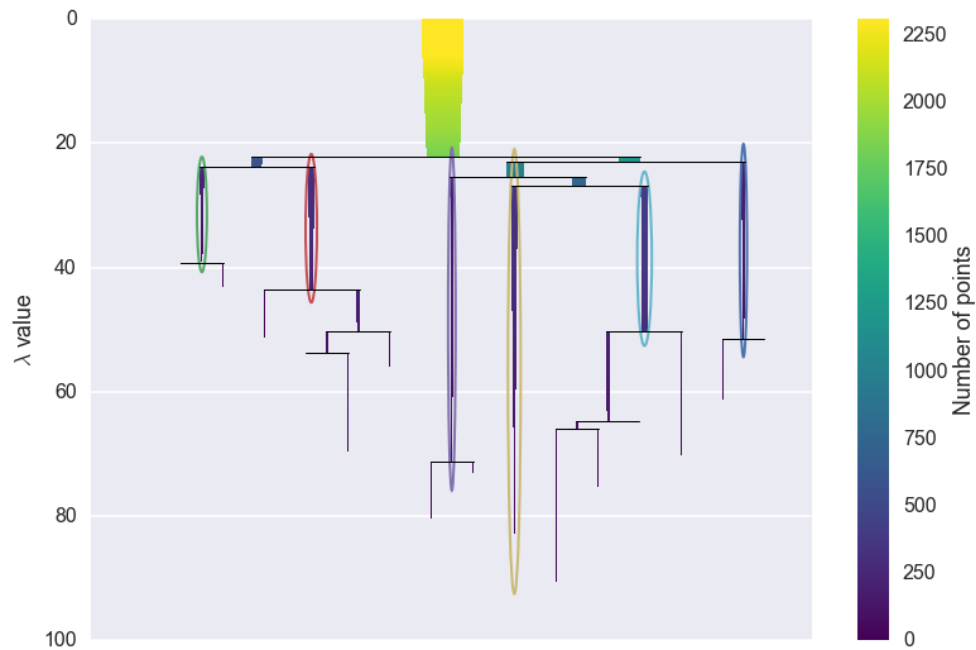
conclusion

Banish J

AI Developer, WEB Developer and APP Developer
to know more about me check <https://www.banish.in/>



Hierarchical DENSITY BASED SPATIAL CLUSTERING APPLICATION



- HDBSCAN

What is HDBSCAN

HDBSCAN - Hierarchical Density-Based Spatial Clustering of Applications with Noise

HDBSCAN is a clustering algorithm that improves on DBSCAN by creating a hierarchy of clusters and doesn't need a fixed radius.

It uses two main parameters:

- 1. Minimum Cluster Size:**

Minimum number of points needed to form a cluster.

- 2. Minimum Samples:**

Minimum number of points to consider a point as dense enough.

Key Concept:

HDBSCAN finds clusters of different shapes and densities without needing a fixed search radius.

1. Core Distance:

The smallest distance at which a point is dense enough.

2. Mutual Reachability Distance:

A measure used to connect points into clusters.

How HDBSCAN Works :

- 1. Calculate Distances:**

Measure how close points are to each other.

- 2. Build a Cluster Tree:**

Create a tree structure showing how points group together at different distances.

- 3. Simplify the Tree:**

Remove less important branches to focus on significant clusters.

- 4. Extract Clusters:**

Identify clusters by finding stable groups in the simplified tree.