GeoDBSCAN

Title GeoDBSCAN

Description

This scripting tool will take in an input feature class and feed each features centroid coordinates into Scikit-Learns DBSCAN Implemenation. This scripting tool will add the resulting DBSCAN Labels directly to the input feature class. DBSCAN requires a neighborhood size, a minimum number of samples per cluster, and an optional weight field to determine clusters.

Source: http://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html

Wiki: https://en.wikipedia.org/wiki/DBSCAN

Application Example: http://geoffboeing.com/2014/08/clustering-to-reduce-spatial-data-set-size/

Usage

DBSCAN - Density-Based Spatial Clustering of Applications with Noise. Is used to find core samples of high density and expands clusters from them. Good for data which contains clusters of similar density. This tool will take points and feed the X-Y coordinates of the current projection as a two dimensional array to be clusted by scikit-learns implementation of DBSCAN.

A cluster satisfies two properties:

All points within the cluster are mutually density-connected.

If a point is density-reachable from any point of the cluster, it is part of the cluster as well.

Source:http://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html

Syntax

GeoDBSCAN (Input_Feature_Class, Neighborhood_Size, Minimum_Samples, {Weight_Field})

Parameter	Explanation	Data Type
Input_Feature_Class	Dialog Reference This is the input feature class that will be clustered using DBSCAN and its labels will be added to the feature class fields. Noise is a -1 and each sequential number after is a cluster. There is no python reference for this parameter.	Feature Layer
Neighborhood_Size	Dialog Reference The maximum distance between two samples for them to be considered as in the same neighborhood. This tool uses a Neighborhood size based on the units of the current projection. The points are represented by the raw centroid coordinates returned by "SHAPE@XY" token.	Double

	There is no python reference for this parameter.	
Minimum_Samples	Dialog Reference The number of samples (or total weight) in a neighborhood for a point to be considered as a core point. This includes the point itself. For example, if you have a feature class that has a weight of 5, and minimum samples is 10, another 5 more points weight 1 or another point of weight 5 to have enough weight to be a cluster. There is no python reference for this parameter.	Long
Weight_Field (Optional)	Dialog Reference Weight of each sample, such that a sample with a weight of at least min_samples is by itself a core sample; a sample with negative weight may inhibit its eps-neighbor from being core. Note that weights are absolute, and default to 1. For each value of the selected field, each point passed There is no python reference for this parameter.	Field

Code Samples

There are no code samples for this tool.

Tags

DBSCAN, Scikit-Learn, Clustering

Credits

David Wasserman

Use limitations

There are no access and use limitations for this item.

You are currently using the Item Description metadata style. Change your metadata style in the Options dialog box to see additional metadata content.