|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **DBSCAN Algorithm: Algorithm to identify and classify anomalies** | | | | |
|  | **Input:** Daily time series (for a given mobile platform if multiple) | | | |
|  | **Output:** DBSCAN-labeled anomalies conceptualized as plumes | | | |
|  | **Initialize** labeledAnoms \\ empty vector = number of total points in mobile TS | | | |
|  | **For** (each daily time series) \\ determine the parameters eps and MinPts and run DBSCAN | | | |
| **1** |  | **Scale** each variable to mean 0 and variance 1 | | |
| **2** |  | **Set** minPts = 0.03 \* n \\ n is the total number of points in the daily mobile monitoring time series | | |
| **3** |  | **Construct** knn ordered distance graph with k = minPts | | |
| **4** |  | **Set** dists = first 30 ordered distances | | |
| **5** |  | **Set** mean = mean of dists | | |
| **6** |  | **Set** sd = standard deviation of dists | | |
| **7** |  | **Set** d = 31st distance in ordered set of distances | | |
| **8** |  | **For** (d, d < total number of distances, d++) \\ Go through remaining distances in the ordered set and find the first distance that is greater than the mean + 3 standard deviations of the set of previous ordered distances | | |
|  |  |  | **If** (d > mean +3 \* sd) | |
|  |  |  |  | **Set** eps = d |
|  |  |  |  | **Break** |
|  |  |  | **Else** \\ Add d to the subset of dists | |
|  |  |  |  | **Concatenate** d to dists |
|  |  |  |  | **Set** mean = mean of dists |
|  |  |  |  | **Set** sd = standard deviation of dists |
|  |  | **End** | | |
|  |  | \\ With eps and MinPts, run DBSCAN on the daily time series | | |
| **9** |  | **Set** dbOutput = dbscan(daily time series, minPts, eps) \\ dbOutput returns DBSCAN labeled core, border, and noise points | | |
| **10** |  | **Set** labeledAnoms = 1 **if** dbOutput is core **else** 2 if dbOutput is border, noise | | |
|  | **End** | | | |