输入：数据集

整数k（代替MinPts和Eps）

1. 遍历数据集，为每个数据对象计算基于信息熵的加权第k距离。

这一步需要用到的：

设数据集D为d维，属性集为，D中数据点在属性上的值记为

数据对象，是第j维数据的权值，则数据对象p,q之间的加权距离为

如果 ，则称属性为离群属性

1. 计算全部数据对象的第k距离的均值
2. 遍历数据集D，对任意，如果满足，则将加入**核心对象**集合
3. 随机选取一个核心对象，以为半径，找出所有可以由核心对象o密度可达的点，生成簇。
4. 跟原始DBSCAN类似，直到所有核心对象均被访问过为止，输出簇划分，以及未被划分进任何簇中的噪声点集合，以及噪声点的个数。
5. 为各个簇中的各个数据对象计算k邻域密度；

并将每个簇中的各个对象的k邻域密度从小到大排列。

1. 在每个簇中，选取k邻域密度较小的一部分的数据对象；

设置比例，，（越大，接下来算法的执行效率越高）即每个簇中选取的数据对象个数为 ;

将以上各个簇中被选取的数据对象及噪声点集合加入新数据集

1. 遍历中的每一个数据对象，计算每个数据对象的可达距离、可达密度、LOF值
2. 将中全部数据对象的LOF值从大到小排序，选择前个数据对象，得到数据集
3. 最终异常数据集合