**蒋文君周总结(20190804)**

本周主要看了DBSCAN、spectral clustering及adaptive neighbor相关的论文，看的比较散，没有目的性，以后主要从基于DTW的矩阵相似性判断、适应性邻居和传递限制时添加概率三个方面进行针对性学习。

**论文1：**RNN-DBSCAN:A Density-Based Clustering algorithm using reverse nearest neignbor density estimates

与传统的DBSCAN不同，RNN-DBSCAN只需要一个参数k，且结合KNN和RNN来求某个点的邻居。此外，RNN-DBSCAN在遍历并标记完所有的点之后，对标记为-1（即噪点，若某个点RNN的个数小于k，则为噪点）的点做一次ExpandClusters()。

**论文2：**Indexing and classifying gigabytes of time series under time warping

本文基于卫星图提取庞大的时间序列数据，为解决此类大规模高维数据问题，提高分类速度，提出了TSI算法，主要结合了基于下限的DTW的层次k-means聚类：1、build tree：在根节点对所有数据集进行k-means操作，生成C，对于C中所有Ci类的数据分别递归做上一步同样的操作，当结点里的数据个数<k，停止递归。2、search tree：设置三个队列dtw\_PQ,lb\_PQ和nn\_PQ,查询的时间序列个数不能超过L。

**论文3：**A Novel Approximation to Dynamic Time Warping allows Anytime Clustering of Massive Time Series Datasets

本文提出了近似DTW的方法，给出上限和下限，上限由欧氏距离求得，下限由LB\_keigh求得，近似DTW结合上限和下限求得。在评价阶段分别应用到hierarchical clustering、k-medoids clustering及spectral clustering观察其效果。（参考了其他论文了解一下这三种聚类）

**论文4：**Septral Embedded Adaptive Neigbors Clustering

本文只有7页，从头到尾都是公式，介绍的非常简洁，并没有看懂是怎么运用adaptive neighbor到septral clustering，打算配合Clustering and projected clustering with adptive neighbors这篇论文去看。

YADING: Fast Clustering of Large-Scale Time Series Data

参考链接：https://blog.csdn.net/qq\_28900249/article/details/84097805

提出算法Yading来提高对大规模时间序列聚类的效率，主要分为三步：1、取样，采取一定的采样策略，在保证采样数据和原数据的数据分布尽可能相似的基础上，取得s个样本；此外，使用逐段聚集平均法对每个样本实例进行降维，降到d维，最终得到采样数据sxd维（原数据NxD维）。2、对采样数据使用𝐿1 norm 相似度衡量及多密度聚类方法进行聚类。3、实施分配策略，将未采样到的属于原数据的数据分配到上一步生成的类中。