

Multi-View Adjacency-Constrained Nearest Neighbor Clustering (Student Abstract)

Jie Yang,¹ Chin-Teng Lin^{1,2}

¹School of Computer Science, University of Technology Sydney, Australia

²Australian AI Institute, University of Technology Sydney, Australia

jie.yang-4@student.uts.edu.au, chin-teng.lin@uts.edu.au

Abstract

Most existing multi-view clustering methods have problems with parameter selection and high computational complexity, and there have been very few works based on hierarchical clustering to learn the complementary information of multiple views. In this paper, we propose a Multi-view Adjacency-constrained Nearest Neighbor Clustering (MANNC) and its parameter-free version (MANNC-PF) to overcome these limitations. Experiments tested on eight real-world datasets validate the superiority of the proposed methods compared with the 13 current state-of-the-art methods.

Introduction

Multi-view clustering has received a lot of attention in recent years. Different from traditional clustering methods, multi-view clustering is exploited to process multi-view data, which is collected from different domains or feature sets. However, the most current approaches suffer from the following two problems in multi-view clustering: a) parameter selection, and b) excessive computational cost. On the other hand, from the perspective of basic clustering principles, many previous multi-view clustering algorithms are based on spectral clustering or subspace clustering, which have some inherent limitations. For example, spectral clustering suffers from the following three problems: a) the instability of results caused by different initializations; b) the K value required to construct adjacency matrix needs to be adjusted; and c) it can only provide clustering results with a single granularity. Few multi-view clustering algorithms are based on hierarchical clustering. Compared with spectral clustering, hierarchical clustering does not require initialization, and a dendrogram can be generated to provide clustering results with different granularity levels.

To solve the issues above, in this paper we propose a Multi-view Adjacency-constrained Nearest Neighbor Clustering algorithm (MANNC). MANNC consists of three main parts: including the Fusion Distance matrices with Extreme Weights (FDEW); Adjacency-constrained Nearest

Neighbor Clustering (ANNC); and the internal evaluation Index based on Rawls' Max-Min criterion (Kameda et al., 2016) (MMI). Given extracting information from each single view is also important (Tao et al., 2018), FDEW aims to learn a fusion distance matrix set, which not only uses complementary information among multiple views, but exploits the information from each single view. ANNC obeys an intuitive rule that one cluster and its nearest neighbor with higher mass (size) should be grouped into the same cluster in the clustering process (Yang and Lin, 2020). ANNC generates multiple partitions based on FDEW. MMI is exploited to choose the best one from the multiple partitions. MANNC just needs to be assigned the desired number of clusters, which can be estimated based on the decision graph of ANNC. Additionally, we propose a parameter-free version of MANNC (MANNC-PF). Without any parameters, MANNC-PF can give partitions at different granularity levels with a time complexity $O(n \log n)$.

Proposed Method

Fusion Distance Matrices with Extreme Weights (FDEW)

Given multi-view data $\{X^{(i)}\}_{i=1}^v$ collected from v views, for i -th view, $X^{(i)} \in R^{n \times dim_i}$, where n and dim_i are the number of data samples and the dimensions of the i -th view respectively. A sample in $X^{(i)}$ can be represented as $x_a^{(i)}$. On one

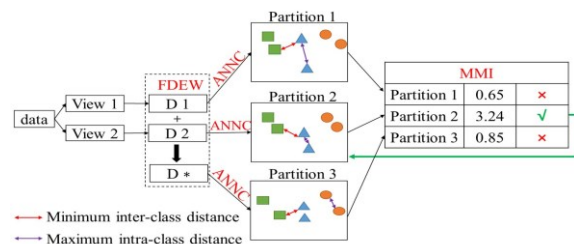


Figure 1. The flowchart of the proposed MANNC.

hand, we regard the cosine distance matrix $D^{(i)}$ of each view as a fusion distance matrix with extreme weights, that is

$$D^{(i)} = 1 \times D^{(i)} + \sum_{j=1, j \neq i}^v 0 \times D^{(j)}, \quad (1)$$

On the other hand, we define a fusion distance matrix with equal weights:

$$D^* = \frac{1}{v} \sum_{i=1}^v D^{(i)}, \quad (2)$$

$D^{(i)}$ only uses the information from each single view, but D^* exploits complementary information among multiple views. Combine $D^{(i)}$ and D^* to form fusion distance matrices with extreme weights $\{FDEW^{(r)}\}_{r=1}^{v+1}$.

Adjacency-constrained Nearest Neighbor Clustering (ANNC) In our previous work (Yang and Lin, 2020), we proposed an adjacency-constrained nearest neighbor clustering (ANNC) algorithm for single-view data.

Given a single-view data $X^{(i)}$, initially, each sample is its own cluster. The following rule is applied to form connections between clusters:

$$\zeta_j \rightarrow \zeta_j^N, \text{ if } \text{mass}(\zeta_j) \leq \text{mass}(\zeta_j^N), \quad (3)$$

where ζ_j denotes the j -th cluster, ζ_j^N denotes the 1-nearest cluster of ζ_j . $\text{mass}(\cdot)$ represents the mass of cluster (i.e., the number of samples contained in the cluster). The symbol " \rightarrow " denotes a connection (i.e., merger) C_j between ζ_j and ζ_j^N . Then, new clusters can be obtained by calculating the connected components of the adjacency matrix. By repeating this merger process according to Eq. (3), all clusters will eventually merge into one cluster and form a hierarchical tree. Each connection (i.e., merger) C_j has two intuitive properties. One of the properties is the product of the mass of the two clusters it connects

$$M_j = \text{mass}(\zeta_j) \times \text{mass}(\zeta_j^N), \quad (4)$$

The other is the square of the distance between the two clusters it connects

$$D_j = d^2(\zeta_j, \zeta_j^N). \quad (5)$$

A decision graph can be constructed by the two properties.

ANNC is parameter-free. A reasonable partition can be obtained through a certain layer (granularity) of the clustering tree, or by observing the decision graph and removing the connections with relatively large M_j and D_j . However, ANNC can also provide a partition with the desired number of clusters K when simply removing $K-1$ connections with

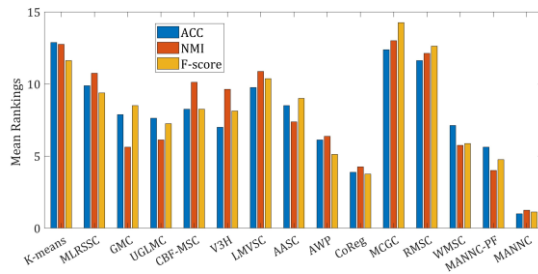


Figure 2. Performance comparison.

largest $M_j \times D_j$. Besides, the complexity of ANNC can be reduced to $O(n \log n)$ via fast approximate nearest neighbor methods, e.g., k-d tree. Exploiting ANNC to perform clustering based on each fusion distance matrix in FDEW, then $v+1$ partitions $\{P^{(r)}\}_{r=1}^{v+1}$ can be obtained.

Internal Evaluation Index based on Rawls' Max-Min Criterion (MMI) Inspired by the max-min criterion (the right decision is that which maximizes the minimum outcome) (Kameda et al., 2016), we propose the internal index MMI to select the best partition from $\{P^{(r)}\}_{r=1}^{v+1}$.

$$MMI^{(r)} = \frac{\min_k \min_l \min_{x_a^{(r)} \in \zeta_k^{(r)}, x_b^{(r)} \in \zeta_l^{(r)}} d(x_a^{(r)}, x_b^{(r)})}{\max_m \frac{2}{|\zeta_m^{(r)}|(|\zeta_m^{(r)}|-1)} \sum_{x_c^{(r)}, x_d^{(r)} \in \zeta_m^{(r)}} d(x_c^{(r)}, x_d^{(r)})}. \quad (6)$$

The partition with the largest MMI is the best partition. Fig. 1 gives the flowchart of the proposed MANNC. Supplementary provides the pseudocode of MANNC and its parameter-free version (MANNC-PF).

Results and Conclusions

We compare MANNC and MANNC-PF with 13 state-of-the-art multi-view clustering algorithms on eight real-world datasets, including 100-leaves, UCI-digits, COIL20, Handwritten, ORL, UMIST, CMU-PIE, and COIL100. Accuracy (ACC), normalized mutual information (NMI), and F-score are exploited to evaluate the performance. Fig. 2 shows the mean rankings for each method on the eight datasets. MANNC ranked No.1 for all evaluation metrics in the mean rankings. Compared with all other methods, MANNC-PF doesn't require any parameters (e.g., number of clusters), but still achieves competitive results.

In conclusion, we propose a novel multi-view clustering algorithm MANNC and its parameter-free version (MANNC-PF). Extensive experiments conducted on eight real-world datasets illustrate their superior performance.

Acknowledgments

This work was supported in part by the Australian Research Council (ARC) under discovery grant DP180100656 and DP210101093.

References

- Kameda, T., Inukai, K., Higuchi, S., Ogawa, A., Kim, H., Matsuda, T., Sakagami, M., 2016. Rawlsian maximin rule operates as a common cognitive anchor in distributive justice and risky decisions. PNAS 113, 11817–11822.
- Tao, H., Hou, C., Liu, X., Liu, T., Yi, D., Zhu, J., 2018. Reliable Multi-View Clustering, in: AAI, 32(1).
- Yang, J., Lin, C.-T., 2020. Clustering via torque balance with mass and distance. arXiv:2004.13160 [cs, stat].