Multi-task Multi-view Clustering using Multi-objective Optimization

Authors

Sayantan Mitra and Sriparna Saha

Department of Computer Science and Engineering,
Indian Institute of Technology Patna, India
Email: sayantaniem@gmail.com & sriparna.saha@gmail.com

Presentation at the 25th International Conference on Pattern Recognition (ICPR2020), 10-15 January, 2021

Overview

- Existing Multi-task Multi-view studies are based on a single-objective function to solve the Multi-task Multi-view learning problems.
- The optimums attained by these methods are often inconsistent.
- Multi-task Multi-view problems are inherently some multi-objective optimization problems because conflict may be between different views within a given task and also between different tasks, necessitating a trade-off.
- In this work, we formulated the multi-task multi-view problem as a multi-objective optimization problem.
- Three objective functions are optimized simultaneously, viz., within-view task relation, within-task view relation and the quality of the clusters obtained.

Architecture

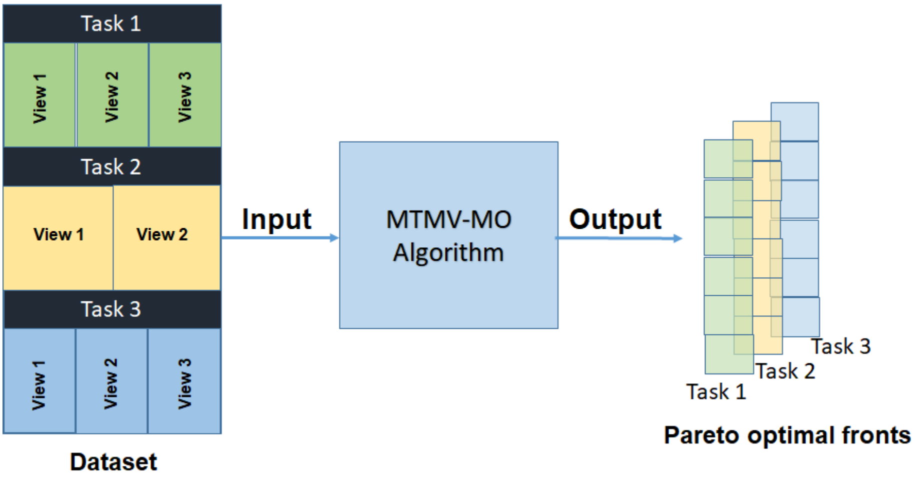

Figure 1: Schematic of the proposed methodology

Flowchart

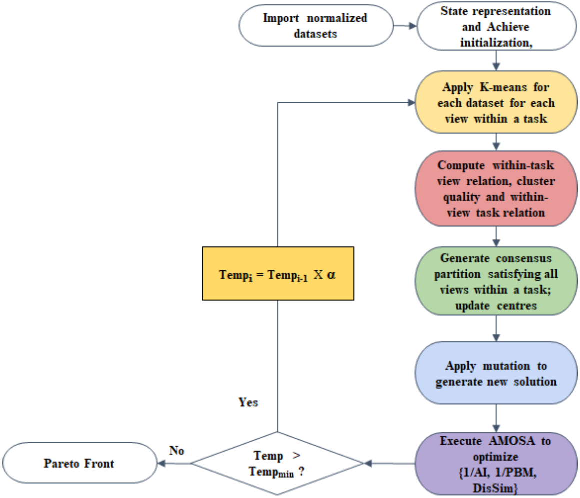

Figure 2: Flowchart of the Algorithm (MTMV-MO)

Within-task View Relation I

- The idea is to maximize the similarity between the clusters among different views for each task.
- For each given task (say t), the Agreement Index (AI_t) between a pair of views (say $v1$ and $v2$) is calculated as follows:

$$AI_{t,v1,v2} = \frac{N_a + 1}{N_d + 1}, \tag{1}$$

$$N_a = \sum_{l=1}^{N} \sum_{m=1}^{N} I_{A_{lm}^{v1}, A_{lm}^{v2}} \tag{2}$$

$$N_d = N^2 - N_a \tag{3}$$

$$I_{A_{lm}^{v1}, A_{lm}^{v2}} = \begin{cases} 1 & \text{if } A_{lm}^{v1} = A_{lm}^{v2} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Here A^{v1} is the adjoint matrix of view $v1$ for the task t; similar for A^{v2}.

Within-task View Relation II

The final *Agreement index* for all the views (V_t) within the task t is given by :

$$AI_t = \frac{\sum_{l=1}^{V_t} \sum_{m=1, l \neq m}^{V_t} 2 \times AI_{v_l l, v_m}}{V_t \times (V_t - 1)} \tag{5}$$

The total *Agreement* index for all the tasks, T, is calculated as:

$$AI = \frac{\sum_{t=1}^{T} AI_t}{T} \tag{6}$$

Cluster Quality Measure I

- To measure the quality of the clusters, we have calculated PBM index [4] for each view within a task; given by PBM_t^v.

$$PBM_t^v = \left(\frac{\mathcal{E}_1 \times D_t^v}{Z \times \mathcal{E}_t^v}\right)$$

$$\mathcal{E}_t^v = \sum_{k=1}^{Z_t^v} \sum_{j=1}^{N_k^v} dist((C_t)_k^v, \overline{n}_j^k)$$

$$D_t^v = \max_{i,j=1}^{Z_t^v} dist((C_t)_i^v, (C_t)_j^v)$$

For a given view v within task t, Z = number of clusters, $(C_t)_i^v$ = center of the i^{th} cluster, $\overline{n}_j^k$ = j^{th} data point of the k^{th} cluster, N_k^v = total number of data points of the k^{th} cluster, D_t^v = maximum distance between a pair of cluster centers and $\mathcal{E}_t^v$ = total distance between all the data points within a cluster and their corresponding cluster centers

Cluster Quality Measure II

- The total PBM index is calculated as the average of all the PBM indices (PBM_t^v).

$$PBM = \frac{\sum_{t=1}^{T} \sum_{v=1}^{V_t} PBM_t^v}{\sum_{t=1}^{T} V_t} \tag{7}$$

Here V_t is the total number of views in the task t.

Within-view Task Relation

- The assumption is that, the clusters formed by the different related tasks for a particular view, should have high similarity.
- If i-th task and j-th tasks are related, then their related cluster centers should have high similarity. This is formulated as follows:

$$dis_{ij} = \sum_{k=1}^{Z} \sum_{l=1}^{Z} \min ||(C_i)_k^v - (C_j)_l^v||_2^2 \tag{8}$$

$$DisSim_v = \sum_{i,j=1, i \neq j}^{T} dis_{ij} \tag{9}$$

The final dissimilarity is calculated by averaging over all the views:

$$DisSim = \frac{\sum_{v=1}^{V_{max}} DisSim_v}{V_{max}} \tag{10}$$

Here V_{max} represents the maximum number of views present. For example, if $Task_1$, $Task_2$ and $Task_3$ have $2, 3$ and 2 views respectively. Then $V_{max} = 3$.
The objective here is to minimize the Eqn. 10.

Dataset Description

Table 1: Description of the datasets.

		Leaves	Mfeat	WebKB	NUS-WiDE
Task 1	Views	3	5	3	7
	Samples	96	1000	226	3615
	Classes	6	5	4	5
Task 2	Views	3	5	3	7
	Samples	96	1000	226	3615
	Classes	6	5	4	5
Task 3	Views	3	-	3	7
	Samples	96	-	226	3615
	Classes	6	-	4	5
Task 4	Views	-	-	3	7
	Samples	-	-	226	3615
	Classes	-	-	4	5

Results

We have compared our proposed methodology with following algorithms:

- **K-Means [2]:** This is a new k-means type algorithm for clustering high-dimensional objects in sub-spaces. It's a simple single-view single-task algorithm.
- **CoRe [3]:** Co-Regularized multi-view spectral clustering algorithm (CoRe) is a spectral based multi-view algorithm.
- **LSSMTC [1]:** This is the shared subspace learning multi-task clustering algorithm (LSSMTC).
- **BMTMVC [5]:** This is the Bipartite graph based multi-task multi-view clustering algorithm (BMTMVC).
- **SMTMVC [5]:** This is the semi-nonnegative matrix tri-factorization based multi-task multi-view clustering algorithm (SMTMVC).

Results

Table 2: Comparison of results in terms of ARI

Methods	Tasks	Leaves	Mfeat	WebKB	NUS-WiDE
K-Means	T 1	61.39	57.16	52.14	46.37
	T 2	58.47	48.61	44.14	40.22
	T 3	59.35	-	41.08	42.52
	T 4	-	-	50.02	43.52
	Avg	90.74	53.14	46.24	43.11
CoRe	T 1	69.11	86.47	60.12	58.34
	T 2	70.13	79.38	59.41	58.12
	T 3	69.92	-	57.44	59.02
	T 4	-	-	60.74	58.41
	Avg	90.72	82.14	60.14	58.42
LSSMTC	T 1	65.98	88.21	83.45	60.25
	T 2	60.56	85.10	81.32	62.53
	T 3	68.12	-	7.14	63.54
	T 4	-	-	82.03	60.02
	Avg	64.24	86.14	81.04	61.51
BMTMVC	T 1	92.54	91.84	90.25	-
	T 2	94.14	85.17	82.84	-
	T 3	93.98	-	85.37	-
	T 4	-	-	86.74	-
	Avg	93.55	89.08	87.14	-
SMTMVC	T 1	92.84	92.08	90.87	67.14
	T 2	94.94	86.18	83.17	68.35
	T 3	94.48	-	86.51	70.14
	T 4	-	-	88.44	64.12
	Avg	94.45	89.18	87.28	66.84
MTMV-MO (proposed)	T 1	93.04	92.05	91.26	68.10
	T 2	95.14	87.10	83.04	69.02
	T 3	96.02	-	86.14	71.04
	T 4	-	-	87.11	65.24
	Avg	94.54	90.85	88.94	67.57

References I

[1] Quanquan Gu and Jie Zhou.
Learning the shared subspace for multi-task clustering and transductive transfer classification.
In *2009 Ninth IEEE International Conference on Data Mining*, pages 159–168. IEEE, 2009.

[2] Liping Jing, Michael K Ng, and Joshua Zhexue Huang.
An entropy weighting k-means algorithm for subspace clustering of high-dimensional sparse data.
IEEE Transactions on Knowledge & Data Engineering, (8):1026–1041, 2007.

[3] Abhishek Kumar and Hal Daumé.
A co-training approach for multi-view spectral clustering.
In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 393–400, 2011.

References II

[4] Malay K Pakhira, Sanghamitra Bandyopadhyay, and Ujjwal Maulik.
Validity index for crisp and fuzzy clusters.
Pattern recognition, 37(3):487–501, 2004.

[5] Xiaotong Zhang, Xianchao Zhang, Han Liu, and Xinyue Liu.
Multi-task multi-view clustering.
IEEE Transactions on Knowledge and Data Engineering, 28(12):3324–3338, 2016.

Acknowledgement

We would like to thank Young Faculty Research Fellowship (YFRF) Award, supported by Visvesvaraya Ph.D. Scheme for Electronics and IT, Ministry of Electronics and Information Technology (MeitY), Government of India, for supporting Dr. Sriparna Saha.

We would also like to extend our sincere gratitude to

- The Anonymous Reviewers
- ICPR 2020 Organizers

Thank You !!

For any further query, please feel free to email me at `sayantaniem@gmail.com`

We wish you all a great virtual ICPR 2020!!