Manifold Adaptive Multiple Kernel K-Means for Clustering

Liang Du  
School of Computer and Information Technology  
Shanxi University  
Taiyuan Shi, Shanxi Sheng, China  
duliang@sxu.edu.cn

Haiying Zhang  
School of Computer and Information Technology  
Shanxi University  
Taiyuan Shi, Shanxi Sheng, China  
1021528684@qq.com

Xin Ren  
School of Computer and Information Technology  
Shanxi University  
Taiyuan Shi, Shanxi Sheng, China  
710372070@qq.com

Xiaolin Lv  
School of Computer and Information Technology  
Shanxi University  
Taiyuan Shi, Shanxi Sheng, China  
969103861@qq.com

ABSTRACT
Multiple kernel methods based on k-means aims to integrate a group of kernels to improve the performance of kernel k-means clustering. However, we observe that most existing multiple kernel k-means methods exploit the nonlinear relationship within kernels, whereas the local manifold structure among multiple kernel space is not sufficiently considered. In this paper, we adopt the manifold adaptive kernel, instead of the original kernel, to integrate the local manifold structure of kernels. Thus, the induced multiple manifold adaptive kernels not only reflect the nonlinear relationship but also the local manifold structure. We then perform multiple kernel clustering within the multiple kernel k-means clustering framework. It has been verified that the proposed method outperforms several state-of-the-art baseline methods on a variety of data sets.

CCS CONCEPTS
• Theory of computation → Kernel methods.

KEYWORDS
Manifold Adaptive Kernel; Multiple Kernel Clustering

ACM Reference Format:
Liang Du, Haiying Zhang, Xin Ren, and Xiaolin Lv. 2020. Manifold Adaptive Multiple Kernel K-Means for Clustering. In ACAI ’20: International Conference on Algorithms, Computing and Artificial Intelligence, December 24–26, 2020, Sanya, China. ACM, Sanya, China, 7 pages. https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

1 INTRODUCTION
Clustering is one of the fundamental topics in data mining, machine learning and pattern recognition. Instead of conducting data clustering within the original feature space, kernel clustering methods perform clustering within the Reproducing Kernel Hilbert space (RKHS), where the nonlinear relationship may be better captured. One of the newly introduced problem for kernel clustering is the design or selection of the proper kernel function, where the best choice is data dependent and unknown in advance.

The Multiple Kernel Clustering (MKC) methods are with great potential to alleviate the effort for kernel designing or integrating complementary information [32] by leveraging a predefined set of candidates kernels from different functions or views. It is natural to extend existing single kernel clustering methods into multiple kernel scenario. The typical methods include K-means based [2, 4, 7, 18, 20, 32, 37, 41], self-organizing map (SOM) [24], maximum margin clustering based [26, 30, 34], local learning-based [33], spectral clustering based [1, 3, 6, 12, 13, 22, 29] and subspace clustering based [8–10, 39, 40] algorithms. Compared with the single kernel counterpart, MKC should take special effort to handle the additional data problems such as noisy and incomplete kernels [2, 15, 19, 21, 28, 36, 38, 40, 41].

Although the candidate kernel well capture the similarities among samples in different nonlinear feature spaces, it dose not necessary characterize the underlying local geometric structure of data, which is vital important for unsupervised learning tasks. Moreover, only a few efforts [14, 27, 39] have been taken to incorporate the local geometric structure of data for MKC. It is worthwhile to point out that these methods use the combined consensus kernel matrix to compute the local graphs, where the discrete neighborhood relationships are changed during the optimization procedure.

In this paper, we present the novel Manifold Adaptive Multiple Kernel K-Means for data clustering (MAMKKC). Given the input kernel matrix, we first construct the corresponding nearest neighbor graph to capture the underlying manifold structure. Then we incorporate the manifold structure into the kernel space via the manifold adaptive kernel mechanism. As a result, the newly induced kernel matrices not only reflect the nonlinear relationship but also the manifold structure. Finally, we linearly combine these manifold adaptive kernels within the multiple kernel k-means clustering framework. We also derive the corresponding optimization procedure to reduce the objective function monotonically and obtain the optimal solution for the proposed MAMKKC model. It has been verified that the proposed method outperforms several state-of-the-art baseline methods on a variety of data sets.
2 MANIFOLD ADAPTIVE MULTIPLE KERNEL K-MEANS

In this paper, we define the number of samples, clusters, kernels as $n$, $c$, $m$, respectively. Suppose that in the clustering task, let $\{K_p\}_{p=1}^m$ denotes the $m$ different kernel functions. correspondingly, there must be $m$ different associated feature spaces denoted as $\{\mathcal{H}_p\}_{p=1}^m$.

The purpose of multiple kernel clustering is to generate the final clustering result via the integration of multiple candidate kernels.

In this paper, we adopt the manifold adaptive kernel transformation [25] to embed the local manifold structure within kernels. Let $\mathcal{V}$ be a linear space with a positive semi-definite inner product (quadratic form) and let $S: \mathcal{H} \rightarrow \mathcal{V}$ be a bounded linear operator. We define $\mathcal{H}$ to be the space of functions from $\mathcal{H}$ with the modified inner product

$$
\langle f, g \rangle_{\mathcal{H}} = \langle f, g \rangle_{\mathcal{H}} + \langle Sf, Sg \rangle_{\mathcal{H}}
$$

(1)

It has been shown that $\mathcal{H}$ is still a RKHS.

Given $f = (f(x_1), \ldots, f(x_n))^T$ and $g = (g(x_1), \ldots, g(x_n))^T$. Notice that $f, g \in \mathcal{V}$, thus we have

$$
\langle Sf, Sg \rangle_{\mathcal{V}} = \langle f, g \rangle = f^T M g.
$$

(2)

where $M$ is a positive semi-definite matrix, and

$$
k_x = (\mathcal{K}(x, x_1), \ldots, \mathcal{K}(x, x_n)).
$$

(3)

It can be shown that the reproducing kernel kernel in $\mathcal{H}$ is

$$
\mathcal{R}(x, z) = \mathcal{H}(x, z) - \lambda k_x^T (1 + \lambda M)^{-1} M k_x
$$

(4)

where $I$ is an identity matrix, $K$ is the kernel matrix in $\mathcal{H}$ and $\lambda \geq 0$ is a constant controlling the smoothness of the functions. 

In this paper, we construct a series of $r$ nearest neighbor graphs $\{G_p\}_{p=1}^m$ to reflect the manifold structure for each candidate kernel. Then, we can construct the associated graph Laplacian $\{L_p\}_{p=1}^m$ via $L_p = D - G_p$. Let $\mathcal{K}$ be any data-independent kernel associated with the kernel matrix $K$. That is, $K_{ij} = \mathcal{K}(x_i, x_j)$. Let $k_i$ be the $i$-th column vector of $K$. By setting $M = L$, we can calculate the manifold adaptive kernel matrix $K_M$ as follows

$$
\mathcal{K}_{M,ij} = K_{ij} - \lambda k_i^T (1 + LK)^{-1} L k_j.
$$

(5)

It is important to note that all the candidate kernels can be transformed to manifold adaptive kernels. In the setting of multiple kernel clustering, we can get many different manifold adaptive kernels, and we linearly integrate these deformed manifold adaptive kernels with different weight.

$$
\mathcal{K}_M = \sum_{p=1}^m w_p^2 \mathcal{K}_p^M
$$

(6)

where $\{w_p\}_{p=1}^m$ is the non-negative kernel weight.

Based on the kernel k-means clustering, we present the novel Manifold Adaptive Multiple Kernel K-Means for data clustering (MAMKKC), which can be formulated as follows

$$
\begin{array}{ll}
\min_{Y, w} & \text{tr} \sum_{p=1}^m w_p^2 \mathcal{K}_p^M (I - YY^T)) \\
\text{s.t.} & w \geq 0, \sum_{p=1}^m w_p = 1, Y^T Y = I, \\
& \mathcal{K}_p^M = K_{ij} - \lambda k_i^T (1 + LK)^{-1} L k_j,
\end{array}
$$

(7)

where tr is the trace function, $\lambda$ is the regularization parameter, $Y$ is the scaled partition matrix, $w$ is the weight of kernels, $I$ is the identity matrix.

3 OPTIMIZATION

As seen from Eq. (7), there are two different variables need to be optimized. We adopted an alternate algorithm which is optimizing one variable while keeping the other fixed.

3.1 Update $Y$

When the $w$ is fixed, the optimization problem with regard to the variable $Y$ can be formulated as follows

$$
\begin{array}{ll}
\max_{Y} & \text{tr}(Y^T N Y) \\
\text{s.t.} & Y^T Y = I,
\end{array}
$$

(8)

where $N = \sum_{p=1}^m w_p^2 \mathcal{K}_p$. The optimal $Y^*$ to the above trace maximization problem can be obtained by the eigenvectors corresponding to the $c$ largest eigenvalues of $N$. The final discrete clustering result then can be obtained via k-means algorithm or spectral rotation from $Y^*$.

3.2 Update $w$

When the variable $Y$ is fixed, the rest optimization problem with regard to the variable $w$ can be formulated as:

$$
\begin{array}{ll}
\min_{w} & w^T A w \\
\text{s.t.} & w \geq 0, \sum_{p=1}^m w_p = 1,
\end{array}
$$

(9)

where $A$ is a diagonal matrix with diagonal element of $A_{pp} = \text{tr}(\mathcal{K}_p - \mathcal{K}_p YY^T)$. The above problem can be solved by off-the-shelf packages.

3.3 Summarization of Algorithm

In sum, we propose the iterative updating algorithm of optimizing Eq. (7) in Algorithm 1. It is obvious that the problem in Eq. (7) is lower bounded. The optimization with respect to $Y$, $w$ will reduce the objective function in Eq. (7) monotonically. Therefore, a local optimal solution can be expected according to our Algorithm 1.

3.4 Algorithm Complexity Analysis

In this subsection, we will discuss the complexity of our proposed algorithm. The computation cost of computing $r$-nearest neighbors of all sample points in all the base kernels, i.e., $\{G_p\}_{p=1}^m$, is
We compared MAMKKC with state-of-the-art multiple kernel clustering algorithms, i.e., LKAMKC [14], ONMKC [20], LKGr [11], JMKSC [31]. It should be pointed that the code for all these 9 methods are obtained from the author’s website or provided by the authors. All the code of our method can also be found at https://gitee.com/csliangdu/MAMKKC, accordingly.

4.1 Compared Algorithms
We compared MAMKKC with state-of-the-art multiple kernel clustering algorithms, i.e., CTSC [12], Coreg [13], RMSC [29], RMKMM [2], LKAMKC [14], ONMKC [20], LKGr [11], JMKSC [31]. In our experiments, our algorithm converges very fast and the times $t$ is less than 20. Since $t < n$, $m < n$ and $\tau < n$, the total computational cost can be simplified as $O(n^3)$.

4 EXPERIMENTS
In this section, we conduct several experiments to evaluate the clustering performance of our proposed algorithm on eight benchmark data sets [5, 16, 17, 35] from various applications, including 3 text corpora ones (BBC, RELATHE, BASEHOCK ), 3 images ones (PIE10P, UMIST, COIL20) and 2 biological ones (Leukemia, ALLAML). The detailed information of these datasets is seen in Table ??.

4.2 Experimental settings
We follow similar settings to prepare 12 base kernels according to [23]. The parameters for these baselines are also set as [23]. For our proposed algorithm, The parameter of proposed algorithm is chosen from $\{1.0, 1.1, \ldots, 2.0\}$. Like previous works, we also use three external measures, i.e., clustering accuracy (ACC), normalized mutual information (NMI) and Purity, to evaluate the clustering results.

4.3 Experimental results
For each clustering algorithm, we report the best results for each parameter corresponding to the best objective values in terms of ACC/NMI/Purity, respectively, from fifty rounds of random initialization in Table 1. We also report the averaged results over all these 8 data sets in the last row of Table 1. It can be seen that MAMKKC consistently outperform other state-of-the-art multiple kernel clustering algorithms. Compared with the second best averaged results, it can be seen that our method achieves 15.75%, 30.79% and 15.92% improvement in terms of ACC/NMI/Purity, respectively. These results show the effectiveness of the proposed method.

For each clustering algorithm, we also calculate the mean ACC/NMI from fifty rounds of random initialization for each parameter and then we additionally report the best mean ACC/NMI together with the standard deviation corresponding to the optimal parameter and the $p$-value of the paired $t$-test against the best results in Table 2, 3. Thus, each cell in Table 2, 3 include the best mean ACC/NMI, the standard deviation and the $p$-value. The best one and those having no significant difference ($p > 0.05$) from the best one are marked in bold. Again, we can observe that our method outperforms better than other MKC algorithms in most cases. And the improvements in most cases are also significant.

For all these compared multiple kernel clustering algorithms, we can observe that the ACC/NMI in Table 1 corresponding to the best objective values are generally higher than the mean ACC/NMI in Table 2, 3.

4.4 Parameter selection and Convergence
For our proposed algorithm, Only one regularization parameter $\lambda$ need to be tuned. As can be seen from Figure 1, it plot the clustering accuracy(ACC) with different values of these parameters on COIL20 and Leukemia respectively. From these figures, it can be seen that the performance of our algorithm is not very sensitive to $\lambda$ within relative wide ranges.

In addition, Figure 2 records the variation trend of the objective function value of our proposed method with increasing number of iterations on three data sets, i.e., COIL20 and Leukemia respectively. As seen from these figures, the objective function is monotonically decreasing, which have also been proofed theoretically. Furthermore, the method quickly converge in less than ten iterations.

**Algorithm 1 The algorithm to solve Eq. (7)**

**Input:** $(K^p)^m_{p=1}, Y, w, \lambda, \tau$
1: repeat
2: Update $K_M$ according to Eq. (6);
3: Update $w$ by solving Eq. (9);
4: Update $Y$ by solving Eq. (8);
5: until Converges
**Output:** $Y, w$

![Figure 1: ACC variants of $\lambda$ on COIL20 and Leukemia](image)

![Figure 2: The convergence on COIL20 and Leukemia](image)
5 CONCLUSIONS AND FUTURE WORK

In this paper, we propose the multiple kernel k-means clustering algorithm with weighted manifold adaptive learning. The proposed MAMKKC algorithm explicitly takes into account the intrinsic manifold structure. The local geometry of the data is captured by a nearest neighbor graph. The graph Laplacian is incorporated into the manifold adaptive kernel space in which multiple kernel clustering is then performed. It can be seen that MAMKKC achieves a good performance compared to many state-of-the-art methods in the extensive experimental.

6 ACKNOWLEDGMENTS

This work is supported in part by the National Natural Science Foundation of China grant 61502289, 61806003, Shanxi Province Key R&D program 201803D31199, Natural Science Foundation of Shanxi Province, China grant No.201810D221163, and Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi STIP 2016101.

REFERENCES

[1] Blake Anderson, Curtis Storlie, and Terran Lane. 2012. Multiple Kernel Learning Clustering with an Application to Malware. In ICMDM. 804–809.
[2] Liang Du, Peng Zhou, Lei Shi, Hammo Wang, Mingyu Fan, Wenhui Wang, and Yi-Dong Shen. 2015. Robust Multiple Kernel K-means Using L21-Norm. In IJCAI. 3476–3482.
[3] Mingyu Fan, Xiaojun Chang, Xiaozing Zhang, Di Wang, and Liang Du. 2017. Top-k Supervise Feature Selection via ADMM for Integer Programming. In IJCAI. 1646–1653.
[4] Mehmet Gonen and Adam A. Margolin. 2014. Localized Data Fusion for Kernel k-Means Clustering with Application to Cancer Biology. In NIPS. 1305–1313.
[5] Nannan Gu, Mingyu Fan, Liang Du, and Dongchun Ren. 2015. Efficient sequential feature selection based on adaptive eigenspace model. Neurocomputing 161 (2015), 199–209.
[6] Hsin-Chien Huang, Yung-Yu Chuang, and Chu-Song Chen. 2012. Affinity aggregation for spectral clustering. In CVPR. 773–780.
[7] Hsin-Chien Huang, Yung-Yu Chuang, and Chu-Song Chen. 2012. Multiple kernel fuzzy clustering. In TAS. 1(2012), 120–134.
[8] Zhao Kang, Chong Peng, and Qiang Cheng. 2017. Kernel-driven similarity learning. Neurocomputing 267 (2017), 210–219.
[9] Zhao Kang, Chong Peng, and Qiang Cheng. 2017. Twin Learning for Similarity and Clustering: A Unified Kernel Approach. In AAAI. 2080–2086.
[10] Zhao Kang, Chong Peng, Qiang Cheng, and Zenglin Xu. 2018. Unified Spectral Clustering With Optimal Graph. In AAAI. 3366–3373.
[11] Zhao Kang, Liangjian Wen, Wenyu Chen, and Zenglin Xu. 2019. Low-rank kernel learning for graph-based clustering. In ICDM. 518–519.
[12] Abyshake Kumar and Hal Daume III. 2011. A Co-training Approach for Multi-view Spectral Clustering. In IJCAI. 393–400.
[13] Abyshake Kumar, Piyush Rai, and Hal Daume III. 2011. Co-regularized Multi-view Spectral Clustering. In NIPS. 1413–1421.
[14] Xiaoliang Li, Xinwang Liu, Lei Wang, Yong Dou, Jianping Yin, and En Zhu. 2016. Multiple Kernel Clustering with Local Kernel Alignment Maximization. In IJCAI. 1704–1710.
[15] Xuan Li, Yang Dou, Xinwang Liu, Yan Zhao, and Qi Lv. 2017. Multiple kernel clustering with corrupted kernels. Neurocomputing 267 (2017), 447–454.
[16] Xinliang Li, Yang Dou, and Liang Du. 2011. Graph-Based Marginal Ranking for Update Summarization. In SDM. 486–497.
[17] Xuan Li, Yi-Dong Shen, Liang Du, and Chen-Yan Xiong. 2010. Exploiting novelty, coverage and balance for topic-focused multi-document summarization. In CIKM.
Table 2: Clustering comparison on (mean ACC)/(standard derivation)/(p-value). The bolder results are significant better than others ($p \leq 0.05$).

| Data Sets | CTSC | Coreg | RMSC | RMKMK | MKKMMR | LKAMKC | ONMKC | LKGr | JMKSC | MAMKKC |
|-----------|------|-------|------|-------|--------|--------|-------|------|-------|--------|
| **BBC**   | 56.18| 54.53 | 55.74| 67.58 | 72.52  | 73.26  | 68.99 | 53.29| 49.83 | 76.62  |
|           | ± 3.83| ± 3.28| ± 3.23| ± 7.80| ± 4.84 | ± 5.14| ± 5.42| ± 3.87| ± 4.62| ± 4.649|
|           | 5.1e-28| 2.7e-34| 3.2e-32| 3.4e-08| 1.5e-04| 6.9e-04| 3.4e-11| 5.4e-33| 1.2e-32| **1.0e+00**|
| **RELATHE**| 64.56| 67.41| 59.73| 56.02| 63.05 | 71.57 | 63.70 | 58.28| 57.18 | 85.35  |
|           | ± 0.29| ± 3.3e-14| ± 2.96| ± 1.07| ± 0.03 | ± 0.14| ± 0.09| ± 0.01| ± 1.0e-13|
|           | 1.1e-92| 0.0e+00| 7.0e-48| 2.4e-72| 6.4e-14| 1.7e-99| 0.0e+00| 1.0e-121| 5.7e-164| **1.0e+00**|
| **PIE10P** | 47.87| 41.23| 36.45| 28.26| 46.27| 53.32| 52.56| 39.39| 54.63| 73.18  |
|           | ± 4.09| ± 3.85| ± 2.76| ± 2.19| ± 4.89| ± 5.34| ± 5.98| ± 4.76| ± 4.76| ± 4.76  |
|           | 2.3e-27| 5.1e-31| 1.9e-34| 5.5e-39| 2.6e-26| 6.8e-21| 2.2e-23| 1.0e-30| 2.4e-20| **1.0e+00**|
| **COIL20** | 62.33| 61.97| 60.46| 61.90| 58.82| 58.97| 59.42| 57.86| 63.07| 72.97  |
|           | ± 3.44| ± 3.80| ± 4.06| ± 3.78| ± 4.43| ± 3.230| ± 4.34| ± 5.83| ± 4.99| ± 4.99  |
|           | 3.5e-18| 9.3e-15| 2.4e-18| 8.1e-18| 7.5e-19| 3.2e-19| 5.8e-21| 7.6e-21| 9.4e-12| **1.0e+00**|
| **UMIST** | 46.17| 44.94| 46.08| 43.36| 47.09| 45.28| 48.59| 43.16| 54.10| 61.72  |
|           | ± 3.35| ± 2.54| ± 3.37| ± 1.84| ± 2.53| ± 2.18| ± 2.65| ± 2.61| ± 4.08| ± 5.10  |
|           | 2.7e-24| 2.1e-26| 5.7e-23| 1.2e-28| 3.9e-24| 3.2e-24| 1.7e-19| 6.4e-28| 8.2e-10| **1.0e+00**|
| **BASEHOCC** | 93.58| 94.76| 94.13| 96.69| 93.10| 95.38| 96.74| 53.84| 52.28| 97.39  |
|           | ± 1.0e-13| ± 0.025| ± 7.8e-14| ± 0.38| ± 0.02| ± 6.7e-14| ± 7.8e-14| ± 4.92| ± 3.3e-14| ± 2.2e-14  |
|           | 0.0e+00| 7.6e-101| 0.0e+00| 3.5e-17| 2.2e-111| 0.0e+00| 0.0e+00| 2.0e-48| 0.0e+00| **1.0e+00**|
| **Leukemia** | 68.60| 60.97| 62.50| 69.11| 63.89| 66.39| 64.39| 70.83| 67.17| 81.94  |
|           | ± 6.7e-14| ± 1.38| ± 0| ± 4.09| ± 3.3e-14| ± 5.87| ± 9.19| ± 2.2e-14| ± 7.09| ± 8.9e-14  |
|           | 0.0e+00| 8.2e-06| 0.0e+00| 3.3e-27| 0.0e+00| 5.5e-24| 3.9e-18| 0.0e+00| 1.2e-19| **1.0e+00**|
| **ALLAML** | 73.61| 66.92| 56.71| 71.83| 66.67| 67.22| 68.56| 76.39| 71.86| 75.03  |
|           | ± 1.1e-14| ± 0.87347| ± 0.6661| ± 3.39| ± 3.3e-14| ± 3.62| ± 6.13| ± 2.2e-14| ± 5.69| ± 12.00  |
|           | 4.0e-01| 1.9e-05| 1.2e-14| **3.8e-02**| 1.0e-05| 8.0e-05| 9.4e-04| **4.2e-01**| 9.7e-02| **1.0e+00**|
| **Average** | 64.04| 61.59| 58.95| 61.84| 63.92| 66.49| 65.37| 56.83| 57.86| 78.03  |
Table 3: Clustering comparison on (mean NMI)/(standard derivation)/(p-value). The bolder results are significant better than others ($p \leq 0.05$).

| Data Sets | CTSC  | Coreg | RMSC  | RMKMKM | MKMMLMK | LKAMKC | ONMKC | LKGr  | JMKSC | MAMKKC |
|-----------|-------|-------|-------|--------|---------|--------|-------|-------|-------|---------|
| BBC       | 38.91 | 37.68 | 37.60 | 50.68  | 45.93   | 50.39  | 47.65 | 38.01 | 24.07 | 59.61   |
|           | ± 1.74| ± 1.97| ± 1.67| ± 6.07 | ± 3.47  | ± 3.56 | ± 4.57| ± 2.35| ± 4.93| ± 3.14  |
| 1.1e-38   | 2.5e-39| 1.6e-39| 3.6e-13| 4.9e-25| 1.8e-17 | 7.5e-21| 1.5e-38| 2.6e-42| 1.0e+00|         |
| RELATHE   | 10.31 | 12.72 | 6.02  | 1.99   | 9.13    | 15.26  | 9.10  | 3.76  | 2.95  | 41.69   |
|           | ± 0.27| ± 5.6e-15| ± 1.11| ± 2.37 | ± 0.02  | ± 0.14 | ± 4.2e-15| ± 0.11| ± 2.1e-15| ± 2.2e-14|
| 1.3e-102  | 0.0e+00| 1.3e-75| 7.8e-62| 3.0e-153| 3.2e-112| 0.0e+00| 8.2e-126| 0.0e+00| 1.0e+00|         |
| PIE10P    | 55.98 | 48.16 | 40.45 | 30.39  | 53.59   | 60.36  | 55.98 | 46.87 | 66.47 | 83.98   |
|           | ± 3.19| ± 3.58| ± 3.14| ± 2.67 | ± 4.42  | ± 4.82 | ± 2.93| ± 3.72| ± 3.76|         |
| 7.2e-38   | 8.0e-43| 5.4e-49| 4.4e-56| 5.6e-36| 7.0e-29 | 1.8e-34| 1.2e-45| 1.6e-28| 1.0e+00|         |
| COIL20    | 75.04 | 75.18 | 74.15 | 73.89  | 74.81   | 71.68  | 72.55 | 69.96 | 78.98 | 86.05   |
|           | ± 1.53| ± 2.07| ± 1.87| ± 2.41 | ± 2.25  | ± 1.69 | ± 2.26| ± 3.39| ± 2.62|         |
| 1.8e-30   | 2.8e-27| 9.0e-30| 2.6e-29| 9.8e-27| 3.1e-33 | 7.3e-32| 3.1e-37| 1.8e-16| 1.0e+00|         |
| UMIST     | 65.07 | 64.80 | 65.35 | 64.13  | 67.23   | 66.63  | 67.36 | 68.61 | 73.09 | 78.91   |
|           | ± 2.2636| ± 1.78| ± 2.77| ± 1.57| ± 1.94  | ± 2.12 | ± 1.78| ± 1.83| ± 1.61| ± 2.41  |
| 3.0e-31   | 4.8e-37| 5.2e-28| 8.0e-37| 2.9e-29| 3.3e-33 | 3.3e-31| 6.8e-31| 4.9e-19| 1.0e+00|         |
| BASEHOCK  | 65.60 | 70.69 | 68.06 | 80.27  | 63.79   | 73.24  | 79.93 | 5.11  | 1.72  | 82.73   |
|           | ± 1.1e-14| ± 0.11| ± 3.3e-14| ± 2.03| ± 0.09  | ± 2.2e-14| ± 4.4e-14| ± 3.68| ± 0.02| ± 1.1e-14|
| 0.0e+00   | 1.9e-100| 0.0e+00| 2.9e-11| 1.3e-114| 0.0e+00| 0.0e+00| 9.2e-67| 2.3e-178| 1.0e+00|         |
| Leukemia  | 9.70  | 5.02  | 5.86  | 11.98  | 6.70    | 7.50   | 9.09  | 12.16 | 22.65 | 26.94   |
|           | ± 2.8e-15| ± 1.35| ± 1.4e-15| ± 2.87| ± 0     | ± 3.77 | ± 7.35| 1.54e-14| ± 4.15| ± 2.8e-14|
| 0.0e+00   | 3.0e-61| 0.0e+00| 2.5e-37| 0.0e+00| 4.0e-37 | 2.3e-22| 0.0e+00| 2.0e-09| 1.0e+00|         |
| ALLAML    | 15.09 | 8.72  | 3.65  | 11.89  | 7.80    | 15.70  | 14.39 | 18.63 | 16.86 | 21.83   |
|           | ± 2.8e-15| ± 0.87| ± 2.09| ± 3.45| ± 3.92  | ± 8.39 | ± 1.27| ± 1.1e-14| ± 2.24| ± 11.52 |
| 1.3e-04   | 1.7e-10| 3.8e-15| 1.7e-07| 3.2e-10| 4.6e-03 | 4.7e-05| 5.5e-02| 4.0e-03| 1.0e+00|         |
| Average   | 41.96 | 40.37 | 37.64 | 40.65  | 41.12   | 45.09  | 44.51 | 32.89 | 35.85 | 60.22   |
