International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

An Analytical Performance Evaluation on Multiview Clustering Approaches

Jyoti R. Mankar[1], Prof. Dr. S. M. Kamalapur[2]
[1]ResearchScholar, Department of Computer Engineering, K. K. Wagh Institute of Engineering Education and Research, Nashik, SavitribaiPhule Pune University, Pune, Maharashtra, India, jrmankar@kkwagh.edu.in
[2] Professor and Research Guide, Department of Computer Engineering, K. K. Wagh Institute of Engineering Education and Research, Nashik, SavitribaiPhule Pune University, Pune,Maharashtra, India , smkamalapur@kkwagh.edu.in

Abstract : The concept of machine learning encompasses a wide variety of different approaches, one of which is called clustering. The data points are grouped together in this approach to the problem. Using a clustering method, it is feasible, given a collection of data points, to classify each data point as belonging to a specific group. This can be done if the algorithm is given the collection of data points. In theory, data points that constitute the same group ought to have attributes and characteristics that are equivalent to one another, however data points that belong to other groups ought to have properties and characteristics that are very different from one another. The generation of multiview data is made possible by recent developments in information collecting technologies. The data were collected from à variety of sources and were analysed using a variety of perspectives. The data in question are what are known as multiview data. On a single view, the conventional clustering algorithms are applied. In spite of this, real-world data are complicated and can be clustered in a variety of different ways, depending on how the data are interpreted. In practise, the real-world data are messy. In recent years, Multiview Clustering, often known as MVC, has garnered an increasing amount of attention due to its goal of utilising complimentary and consensus information derived from different points of view. On the other hand, the vast majority of the systems that are currently available only enable the single-clustering scenario, whereby only makes utilization of a single cluster to split the data. This is the case since there is only one cluster accessible. In light of this, it is absolutely necessary to carry out investigation on the multiview data format. The study work is centred on multiview clustering and how well it performs compared to these other strategies.

Keywords: Multi-Kernel Learning, Co-Training, Graph Clustering, Multi-Task Learning, Multiview Clustering; Non-Negative Matrix Factorization, Subspace Clustering.

I. INTRODUCTION

Clustering is one of the most important unsupervised learning techniques, and it has found a variety of applications in the field of data analysis, including social network analysis, gene expression analysis, heterogeneous data analysis, and market analysis. Clustering [1] is one of the most important unsupervised learning techniques. The purpose of clustering is to divide a data set into many groups in such a way that the data samples included within each group are more similar to one another than the data samples contained throughout the different groups. The process of mining the hidden patterns relies heavily on clustering to organise the data.The data that are taken from the real world always come from a variety of sources and are always represented by a number of different featureviews. Various perspectives on the data each describe a unique subset of the data's characteristics. If you cluster based on information that is complementary to each other and is provided by several views, you will achieve better results than if you cluster based just on one view.Historically, it was believed that one particular subset was all that was required for data mining, and many perspectives were frequently considered to be unnecessary duplications of the same information. But recent studies have shown that looking at data from a variety of perspectives can often yield complimentary results and contribute to a deeper comprehension of the underlying data structure. Multi view learning [4][5] has two advantages: first, a better performance can be obtained by integrating the many views rather than using a single view, and second, the correctness of the knowledge that is created has the potential to be cross-verified through multiple perspectives. As a result, multi-view clustering is required in order to effectively manage multiview data.

In the realm of machine learning, the discipline of multiview clustering (MVC) [3], along with many other multi-view or multi-modal subfields, is predicated on the concept that not only the addition of additional data, but also the incorporation of various kinds of data, can result in improved outcomes. The parable of "the blind men and the elephant" [2] is a helpful illustration for comprehending the significance of MVC, also known as multiview learning. In this allegory, each blind man stands for a different perspective on the matter at hand. However, given that no one perspective on the subject can ever obtain a whole picture of the subject, the only way to ever recover an entire picture of the subject is to compile the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

information obtained from a variety of different perspectives. To provide further clarity, each perspective represents one different source of information. For instance, the contents of a web page can be used to describe the page from one perspective, but the information included in hyperlinks can be seen from a different perspective. Additionally, distinct aspects of a datum can be considered to be diverse viewpoints of the same thing. In the field of image processing, every image is broken down and analyzed using a variety of features, including CENTRIST, Color Moment, Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Scale Invariant and Feature Transformation (SIFT). CENTRIST is a feature that measures the intensity of a single color across time. As shown in figure 1, Original image is (a) and its representation using 5 views are shown in (b),(c),(d),(e)and (f). Similarly original image (g) and its representation using 5 views are shown in (h),(i),(j),(k)and (l) and original image (m) and its representation using 5 views are shown in (n),(o),(p),(q)and (r). Multi-view algorithms deal with each view of the data individually and then combine the solutions to get a comprehensive, robust pattern that is superior to its single-view representation. This is because multi-view algorithms deal with the data in a way that is more like how humans think about it.

It is important to note, for the purpose of outlining two illustrative domains, that health care facilities frequently capture the same disease condition utilizing different medical sensors (for example, EEG, fMRI, and PET are a variety of methods of capturing neurological in nature information), and that criminal history records frequently represent the same crime utilizing techniques such as textual narratives, CCTV footages, audio tapes, and photographs. The purpose of this is to illustrate the similarities and differences that exist between the two types of domains. Image categorization, pattern recognition, and motion segmentation are just some of the applications that have found widespread use for multiview clustering.

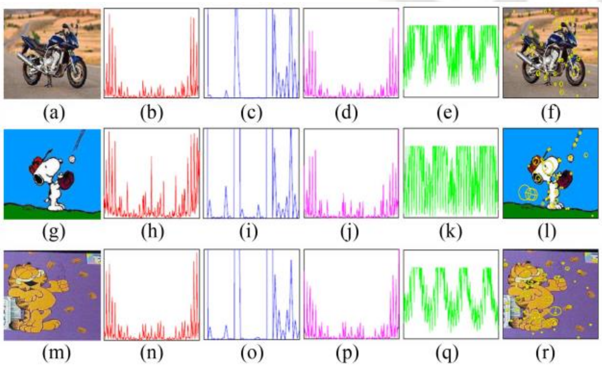

Figure 1: Examples of multiview data.

The investigation of multiview clustering methods is the most important addition that this research makes. The first section of the paper is all about the different multiview clustering methods. It is indicated in Section II that there exist multiview benchmarking datasets. Performance analysis makes up Section III of this document.

II. MULTIVIEW CLUSTERING APPROACHES

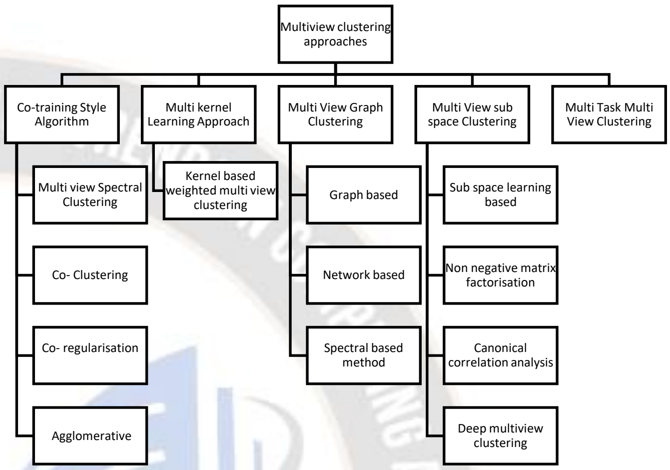

Figure 2: Multiview clustering approaches

This literature review is comprised of a variety of multiview clustering approaches that are representative of the field. It is broken down and summarised in five different categories, as shown in figure 2, in accordance with the mechanisms and principles that form the basis of these methodologies.

1. Co-training style algorithms:

Co-training is a strategy for machine learning that is utilized in scenarios in which here only exists small amounts of labelled data but significant volumes of unlabeled data. In these kinds of scenarios, the unlabeled data is more important than the labelled data. Text mining, which is used by search engines, is one of its applications. It does this by operating under the presumption that every instance can be characterized by employing two different sets of attributes, each of which offers information that is complimentary to the other concerning the instance. Each perspective is sufficient (that is, the class of an instance can be accurately predicted from each view alone), and each of the perspectives have integration (the objective functions export the same predictions for co-occurring features with a significant likelihood of occurring in both views). This ideal scenario occurs when the two perceives are conditionally detached from one another (that is, the two feature sets of each instance are conditionally independent given the class). Bickel and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

Scheffer[6] originally examined MVC with a co-training approach in the context of unsupervised learning (also known as clustering), and they suggested two different forms of MVC algorithms for use with text data. Expectation maximisation (EM) [7] algorithm, which works in an iterative manner between the views, and an agglomerative algorithm [8] that is influenced by the co-training algorithm.

This class of approaches seeks to achieve the greatest possible degree of concordance among all points of view and arrive at the widest possible consensus. Figure 3 illustrates the usual co-training algorithms' standard operating procedure for the training process.Utilising previous knowledge or gaining new information from one another is how it gets the process of clustering various perspectives started. When this tactic is applied in an iterative manner, the outcomes of the clustering of all perspectives become likely to one another, which ultimately results in the widest possible consensus across all perspectives. The clustering findings from one view are utilized for restricting the level of resemblance for the other views, which is the fundamental premise upon which the co-training techniques are founded. This is the fundamental idea behind these algorithms.

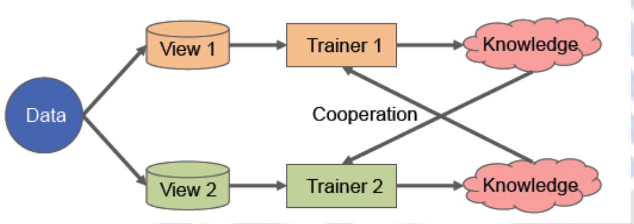

Figure 3: General procedure of co-training

1.1 Co- regularization

In order to bootstrap classifiers in each view, the Co-training approach uses unlabeled samples from many perspectives, often in a greedy manner. It also operates under the notion that perspectives are compatible with one another and are independent of one another. Regularisation is a technique that is used to cut down on errors by adequately fitting the function on the provided training set and avoiding overfitting at the same time. In a paradigm known as Co-Regularization [9], classifiers are learned in each view using various forms of multi-view regularisation. This framework for co-regularization is founded on the principle of optimising metrics of agreement and smoothness over samples that have either been labelled or not. The Co-Regularized Least Squares (Co-RLS) [10] algorithm is part of the family of algorithms that make up this framework. This algorithm does a joint regularization that aims to minimize disagreement in a least squared sense. In addition, the Co-Regularized Laplacian Support Vector Machine (SVM) [11] and Least Squares (Co-LapSVM, Co-LapRLS) techniques employ multiview graph regularises to impose complementary and robust ideas of smoothness in each view.

Some salient features of the this algorithms are

- The traditional framework of regularisation in replicating Kernel Hilbert Spaces is extended naturally, giving rise to these techniques. Additional regularizers that are driven from established semi-supervised learning concepts are utilised in order to incorporate the unlabeled data.
- The algorithms are not greedy; they utilise convex cost functions; and it is simple to put them into action.
- It is possible to exercise explicit control over the influence of both numerous views and unlabeled data. The fundamental concept behind co-regularization is to reduce the amount of difference that exists between the predictor functions of two different points of view, with this reduction serving as an element of the objective function.

1.2 Multi view Spectral Clustering

The difficulty of finding the minimal cut in graphs is intricately connected to spectral clustering. To begin, it performs dimensionality reduction on the initial data space by utilizing the spectrum of the similarity matrix of data samples. This is done in order to simplify the analysis. After that, it applies the k-means clustering algorithm to the low-dimensional space in order to divide the data into a number of distinct groups. For this reason, the initial step in the process of analysing a group of data samples should involve the construction of a similarity matrix. The goal of spectral clustering is to organise data into distinct groups by efficiently investigating complimentary information from a number of different Laplacian matrices. Spectral clustering is a method that has its origins in graph theory. In this field of study, the method is utilised to determine communities of nodes in a network based on the edges that connect them. The method is adaptable and supports the clustering of data that does not contain graphs. When performing spectral clustering, the information required comes from the spectrum, or eigen values, of specialised matrices that are constructed from either the graph or the data collection.

Steps for co-training based multiview spectral clustering:

1. Solve the spectral clustering problem using each graph individually to obtain the discriminative eigenvectors for each perspective, such as U1 and U2.
2. Cluster the points with the help of U1, and then utilize this clustering to change the structure of the graph in view 2.
3. Cluster the points using the U2 algorithm, and then utilize this clustering to make changes to the network structure in view 1.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

4. Return to Step 1 and go on doing so for the desired number of times.

Better clustering results have been obtained by a number of earlier approaches [12,13,14] that were related to co-training. These techniques all use spectral clustering, which is advantageous for graph clustering because it can accept irregular cluster shapes and is hence more efficient. Co-training is a technique to multi-view spectral clustering that was proposed by Kumar and Daum eIII[12]. In this method, the clusters are bootstrapped to different perspectives by utilising complementing information from one another. The most recent method of multi-view clustering, which is based on co-training, employs the utilisation of spectral embedding from one view to update other views. In addition, Kumar et al. proposed a co-regularized technique for multi-view spectral clustering in their paper [13]. In this method, the graph Laplacians are put in place on all views, and regularizations are performed on the Eigenvectors are of the Laplacians. This is done in order to guide the resulting clustering arrangements in a consistent manner.

Because this particular low-dimensional encoding is the one that is used to generate clusters, the spectral embedding of a view can be regarded of as representing an underlying clustering structure. This is because this particular low-dimensional embedding is the one that is used to build clusters. However, the spectral embedding does not reflect unambiguous borders between clusters; as a result, it has an impact on the ability of the co-training algorithm to converge.

Sally El Hajjar and Fadi Dornaika's paper [15] outlines an innovative method for doing graph-based multiview clustering in one step. Two significant advancements are included, both of which set it apart from other graph-based approaches that cluster in a single step. In the beginning, construct an additional graph by applying the cluster label correlation to the graphs that are connected to the data space. Second, a smoothing constraint is utilised in order to compress the cluster-label matrix thereby rendering it more consistent with the initial data graphs as well as with and label graphs. This is accomplished through the utilisation of a smoothing constraint. In addition, the spectral clustering technique that is used in the procedure of clustering has problems with both its scalability and its runtime.

1.3 Co- Clustering

There has also been research done on multiview clustering that is based on co-clustering, which involves simultaneously clustering the objects and the characteristics. Clustering challenge posed by a data matrix with features (variables, or attributes) organised in columns and objects (attributes, or rows) in the matrix's rows. The most fundamental method of data mining is known as clustering items based on the data matrix. This method groups objects that have similar patterns of distribution. Co-clustering is an extension of traditional clustering that involves the creation of a model that not only captures the structure of object clusters but also the structure of feature clusters.

For instance, Meng et al. [16] presented a heterogeneous data co-clustering technique. This strategy not only broadens the scope of fusion from two views to many views, but it also assigns different data sources different weights based on the features they contribute. This approach clusters heterogeneous data in a way that allows for multiple views of the same data. The approximate alternate linearized minimization strategy was presented by Sun et al. [17]. The matrix decomposition served as the foundation for this approach.Using this strategy, it is possible to deconstruct many data matrix structures into sparse row and column vectors at the exact same time. In addition to this, it is able to link the various views of the data by making use of a binary vector. This binary vector ensures that the row clusters remain consistent throughout all of the views. In [18], the architecture for learning co-similarities from multi-view datasets was designed. In [19], the architecture was then parallelized. The architecture was designed to simultaneously build similarity matrices between the rows and columns of a data matrix, rather than a set of clusters; this was done in order to learn co-similarities. In addition to this, it is able to link the various views of the data by making use of a binary vector. This binary vector ensures that the row clusters remain consistent throughout all of the views. In [18], the architecture for learning co-similarities using multi-view datasets was designed. In [19], the architecture was then parallelized. They made this assumption based on the fact that they believed that transferring similarity values that were generated from individual data from one view to others would result in improved data clustering. In addition, a number of collaborative Multiview clustering strategies have been looked into in [21, 22]. The first step is conducted locally, while the second stage is conducted in partnership with others. A clustering method is applied to each view during the local phase, and then, during the collaboration phase, each view is coupled with the clustering outcomes corresponding to the other views that have been generated during the local phase. During the local phase, each view is clustered independently of the other views.

The purpose of performing multiple clusterings is to identify a number of different approaches that can be used to independently organize a dataset into clusters. The current solutions to this problem only concentrate on clustering in a single direction. However, in many applications that are used in the real world, it is useful and desired to investigate alternative

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

two-way clustering (also known as co-clusterings) [18]. This is the process of clustering both samples and characteristics.

By enabling the automatic discovery of similarity based on a subset of qualities, it is able to circumvent a number of the restrictions that are inherent to conventional clustering algorithms.

A user's rating of a document, for instance, is influenced not just by the user's qualities (such as their affinity for certain subjects or categories), but also by the document's features (such as its connections to one or more of those subjects or categories).

The performance of co-clustering is improved for datasets that are both sparse and tiny. If the data are positive, then it works very well.

1.4 Agglomerative

The techniques used in agglomerative clustering work by repeatedly merging the clusters that are geographically closest to one other. A obvious extension of this process for the multi-view context is to divide up the iterative merging procedure in such a way that one iteration runs one merging step in a particular view and the next iteration step in the other view and so on. This extension of the procedure is called "splitting up the iterative merging procedure."

An example of a hierarchical clustering algorithm is referred to as agglomerative clustering. The bottom-up methodology, which is used to organize the datasets into clusters, is shown in figure 4, and can be seen as an example of this methodology. This means that the approach begins by treating each dataset as a component of a single cluster, and then it begins merging the clusters that are the most closely associated with one another. In the end, it creates a new cluster that contains all of the datasets that are the most closely connected to one another.

The step that Agglomerative Clustering take are:

1. Determine the distance measurement, and then compute the distance matrix, with each data point being assigned to a particular cluster.
2. Find out the conditions for linking so that the clusters can be merged.
3. Keep the distance matrix up to date.
4. Continue the process until each data point is represented by a single cluster.

Bickel and Scheffer [6] initially researched MVC with the idea of co-training, and they suggested two different forms of Multiviewclustering methods for text data. The first is an EM algorithm with multiple perspectives that operates by switching back and forth between the views, and the second is an agglomerative algorithm that is modeled after the co-training method. As a consequence of this, Bickel and Scheffer [6] came to the realization that the multiview EM algorithm performed noticeably better than the single-view approach. On the other hand, the agglomerative algorithm produced unfavorable outcomes.

2. Multi kernel multiview clustering

The term "multiple kernel learning" refers to a group of different techniques to machine learning that make use of a predetermined set of kernels and learn the optimum linear or non-linear combination of one or more kernels as an integrated component of the algorithm. These approaches make use of a given set of kernels in order to accomplish their goals. These methods make use of a collection of kernels that has already been defined and learn the most efficient linear or non-linear combinations of kernels that produce the best results. These methods enable the data to be categorized in a wide range of different ways. Learning from multiple kernels can be helpful for a variety of reasons, including the ones that are listed below: a) the ability to select for an optimal kernel and the parameters from a larger set of kernels, thereby reducing bias due to kernel selection while simultaneously allowing for more automated machine learning methods; b) the ability to bring together data from numerous sources (for example, sound and images from a video) that possess distinct notions of similarity and, as a result, require different kernels; both of these advantages are related to each other and are important to understand. It is possible to integrate kernels that have already been produced for each separate data source by employing several kernel methods rather than developing a whole new kernel. In a scenario with several viewpoints, De Sa et al. [23] developed a specialized kernel combination approach that was based on an algorithm that minimized the amount of disagreement [24, 25]. To be more specific, they constructed a multi-partite graph with the goal of inducing a kernel, which was subsequently put to use for spectral clustering. This method can be considered an alternative to kernel canonical correlation evaluation, as well as an expanded version of co-clustering and spectral clustering. Additionally, it can be viewed as a clustering methodology that combines the two. In addition, Yu et al. [26] incorporated Hilbert space into the conventional K-means clustering. The multi-view data matrices used inside this new framework were referred to as kernel matrices, and they were automatically integrated for the purpose of data fusion [27]. The kernels were combined in a way that was unique to the region in question in order to provide a more accurate representation of the sample features included within the data [28]. Lu et al. [29] focused their research on multiple kernel grouping based on a centered kernel alignment, which is an efficient kernel evaluation metric. This was the subject of the research that they did. This

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
Article Received: 28 April 2023 Revised: 10 June 2023 Accepted: 27 June 2023

was done so that a single optimization framework could be created out of two different clustering tasks as well as multi-kernel learning. Research has also been done on methods that use a weighted combination of kernels, and this research takes into account the differences in perspectives (or kernels). For instance, the weights that were assigned to the kernels were determined by the information quality of the views that corresponded to those kernels [30]. The kernel proper alignment served as the foundation for the learning of the kernel matrix's structure, which was done in order to determine the degree of similarity between two kernel matrices. In addition, Liu et al. [30] demonstrated a weighted numerous kernel K-means clustering algorithm that made use of matrix-induced normalization of employment. The above technique was able to automatically determine weights for deriving the kernel matrix on each view through an optimization process.

The multi-kernel MVC architecture is more difficult to grasp and requires more processing power.

In numerous applications, including as event detection in video, object recognition in pictures, and biological data fusion, multiple kernel learning strategies have been utilized.

2.1 Kernel based weighted multiview clustering

Views are described in terms of the kernel matrices that are provided in this approach, and a weighted combination of the kernels is learned in parallel with the partitioning process. The quality of the information included in the views that correspond to the kernels is reflected in the weights that are assigned to the kernels. In addition to this, the combination scheme includes a parameter that regulates the permissible sparsity of the weights. This helps to prevent extremes and ensures that the weights are appropriate for the data. In this paper, two efficient iterative techniques are shown that maximize the intra-cluster variance from distinct points of view by alternating between upgrading the view weights and recomputing the clusters.

A weighted multiview clustering approach that incorporates matrix-induced and low rank regularization was presented by Zhao et al. [32]. Zhang et al. [33] also created an auto-weighted multi-kernel multiview clustering technique that simultaneously weights the views and kernels. Both of these algorithms were presented as part of a weighted multiview clustering approach that was based on modified Gaussian kernels that had changeable weights. Trivedi et al. [34] proposed a general technique that permits the multi view the case of clustering in complete view situations, to be applicable in this scenario, where just one perspective was complete while the auxiliary views were partial. This technique enables the multiview clustering to be applicable in completeview circumstances since it is a generic technique. Within the scope of their paper, this methodology was discussed. In order to demonstrate their point, they used a method called kernel canonical correlation analysis (CCA) based multiview clustering. According to De Sa et al. [23], their suggested approach was able to assess sample affinities even when there were missing views. Shao et al. [35] created a collective kernel learning technique as a means of deducing the hidden sample similarity in a scenario in which no single perspective can be regarded comprehensive. This technique was designed for use in situations like the one described above. This strategy's objective is to maximize the alignment of the shared instances of those views in order to comprehensively finish the kernel matrices of the unfinished views as a group. In addition, Liu et al. [36] merged the kernel imputation and clustering into a unified learning technique for imperfect multiview clustering. This is in contrast to other existing methods, which do not include this integration. In those other methods, entire kernels were initially guessed, and then a multi-kernel clustering procedure that was readily available was utilized on the kernels that were input. This was done in order to cluster the input kernels.

This Multi kernel weighted multiview clustering method not worked well for text data.

3. Multiview graph clustering

Each data object is represented by a node in the graph, and the connection that exists between each pair of objects is represented by an edge between those nodes. Graphs, also known as networks, are frequently utilized for the purpose of describing the relationships that exist between objects. In point of fact, the similarity or the affinity connection is the one that is used to determine the nature of the connection most of the time; more particularly, the input graph matrix is constructed from an information similarity matrix. Data objects are represented by numerous graphs at the same time in a multiview situation. It is a widely held belief that any unique graph can only represent a subset of the information contained in the data, despite the fact that all graphs share the same underlying data clustering structure. As a consequence of this, these graphs are able to mutually reinforce one another by collectively consolidating the correlation that exists between the various data elements. The graph-based fusion technique for multiview data is, in general, very similar to figure 4.The objective of learning low-dimensional representations of nodes in a multi-view network using multi-view graph embedding is to capture the many relationships between the nodes. Every one of the views that make up the multi-view network depicts a different manner in which the nodes interact with one another. This series of procedures begins by looking for a fusion graph (or network) across all views. After that, graph-cut techniques or different techniques (such as spectral clustering, for example) are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
Article Received: 28 April 2023 Revised: 10 June 2023 Accepted: 27 June 2023

performed on the fusion network in order to achieve a clustering result.

3.1 Graph basedmultiviewclustering

The extraction of common elements from numerous graphs through the use of certain methods, such as the linked matrices factorization method, led to the creation of a variety of graph-based algorithmic clustering techniques that were easily applicable to multiview data. These methods are based on the concept of multiple similarity graphs. An technique for clustering documents using multiple viewpoints was proposed by Hussain et al. [37]. In order to build numerous partitions, the single-view clustering methods are performed to the data matrices of each view. After that, these partitions provide a set of three separate similarity matrices, which are respectively known as the affinity matrix, the cluster-based similarity matrix, and the pair-wise dissimilarity matrix. Each of these matrices is named after the clusters that they are based on. In the end, it uses a method known as an ensemble to aggregate all of these matrix structures into a single, unified similarity matrix that can then be used for clustering. This matrix may then be used to group items together. In addition, the impact of various similarity metrics (such as Pearson and Spearman correlations, Euclidean and Canberra Distances, etc.) on multiview clustering has been investigated in [56]. This research may be found in the citation below. Xue et al. [57] presented a group-aware multi-view fusion methodology as their contribution. This approach makes use of a variety of weights in order to quantify the degree of pair wise similarity that exists between diverse groupings. Even though certain multiview clustering strategies were able to acquire a weight for each graph, such systems still contain additional parameters to adjust. Nieet al. [58] developed a framework for numerous graphs that is free of parameters and can automatically learn a set of weights to apply to each graph in order to address the challenges described above. Additionally, unsupervised feature selection for multi-view data is a technique that has also been investigated as part of this project. These particular qualities were selected for use in a clustering activity in addition to other learning activities since they lend themselves well to such applications.

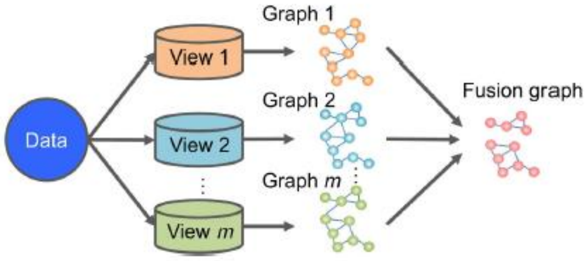

Figure 4: General procedure of graph-based clustering

3.2 Network based multiview clustering

The vast majority of graph-based MVC techniques generally make the assumption that the same collection of data items is accessible for each view. As a result, the relationship that exists between data items in the various views is a one-to-one relationship. A many-to-many mapping relationship is created when an object in one domain can correspond to several objects in another domain. This type of relationship is also known as a one-to-many mapping relationship. This is the situation in a number of various programs that have their foundations on real life, which include social networks, literature citation networks, and biological interaction networks. It's possible that depicting these connections using networks, rather than graphs, might be a more accurate representation. The primary difference between network-based MVC and graph-based MVC is due to this particular feature.

To put it another manner, an object that is present in one domain could correlate to multiple other objects that are present in a different domain.

In addition, unique data distributions are characteristic of each individual domain. This undermines a second assumption, which held that all perspectives would have the same grouping structure. Ni et al. [38] created a robust and adaptable multi network clustering framework to overcome the aforementioned two difficulties. This framework permits many-to-many relationships and many underlying clustering structures.

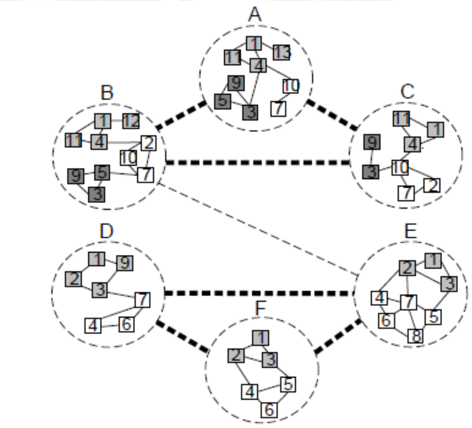

Figure 5: Network of network

In particular, Ni et al. [38] approached the degree of similarity among each domain as if it were its own network. Additionally, they modeled the similarity across the multiple domains as if it were its own network in order to normalize the clustering structures that were present in the various networks. The definition of a global network, which has been referred to as a Network of Networks (NoN), is depicted in the figure 5 that can be found attached to this paragraph. In this picture, the network that is dashed symbolizes the main network that

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

connects the six different domains A, B, C, D, E, and F, and each node in the main network corresponds to a domain-specific network that is represented by the solid lines. Main clusterings that have and domain clusterings are names given to the groupings of nodes that exist in the primary network and the domain-specific network, respectively. These names are meant to be used in conjunction with one another. In light of these ideas, they settled on a strategy consisting of two stages in order to partition the NoN. This strategy begins with the partitioning of the main network, which is followed by the incorporation of the learning information obtained from the main network into the clustering of domain-specific networks.

3.3 Spectral based multiview clustering

The data clustering paradigm known as spectral clustering has been around for a long time. Forming a pair-wise affinity matrix between any two objects, then normalizing this affinity matrix, and finally computing the eigenvectors of this normalized affinity matrix (also known as the graph Laplacian) are the fundamental steps in this process. It has been shown that the subsequent eigenvector of the normalized graph Laplacian is a compression of a binary vector solution. This finding was made possible by the fact that the graph was normalized. This conclusion was reached as a result of the previous sentence. This approach has the potential to reduce the normalized cut on a graph, which refers to the connection between the spectral and the graph. De Sa devised a spectrum clustering approach on two independent perspectives in his paper [39], and each of these views could be fed into a separate clustering model. In order to connect the two-view features, this spectral-based multiview clustering method first generated a bipartite graph using a minimizing-disagreement criterion [40, 41]. On this bipartite graph, it proceeded to execute a number of different spectral clustering techniques. Zhou and Burges [42] investigated multiview spectral clustering by extrapolating a well-known cut from the single view to multiple views; examining how to identify a clustering algorithm that is close to the best possible result for all graphs; and further developing a multiview transductive deductive approach on the basis of multiview spectral clustering. This was done. All of these were done in order to further develop a multiview transductive deductive reasoning on the basis of multiview spectral clustering. Zhou and Burges [42] also investigated how to learn a clustering close to the optimal solution for all graphs. Work of a similar nature has been done [43], with the intention of locating a balance cut that is capable of effectively separating all similarity graphs from one another.

4. Multiview subspace clustering

The process of substructure clustering is an extension of classic clustering that aims to locate clusters in a dataset's various subspaces in addition to the original dataset's whole. When dealing with high-dimensional data, it is common for many of the dimensions to be irrelevant, which might disguise existent clusters in data that is noisy. Through careful examination of the full dataset, the process of feature selection eliminates dimensions that are deemed unnecessary or duplicated. The search for relevant dimensions is localized by the algorithms used for subspace clustering, which enables them to locate clusters that are present in several subspaces and may overlap with one another. The problem of multi-view subspace the case of clustering which refers to the process of developing a new and unified depictions for all view data, from multiple subspaces, or a space of latent information that makes it easier to cope with highly dimensional information when creating clustering models, has recently been an important subject in the field of multiview clustering. This problem pertains to the process of acquiring new and unified visualizations for each and every set of view data, which originates from a number of different subspaces. Specifically, this procedure requires learning a new representation that is unified for all of the view data that comes from the various subspaces.

Figure 6 provides a visual representation of the multiview subspace clustering process in its overall form. One can obtain a unified representation of the features in one of two ways:

I. Discover a unified representation by directly learning from many subspaces.
II. Discovering a hidden space is the first step toward reaching this overarching representation. In the end, this unified description was used to obtain the clustering findings by feeding it into a pre-made clustering model.

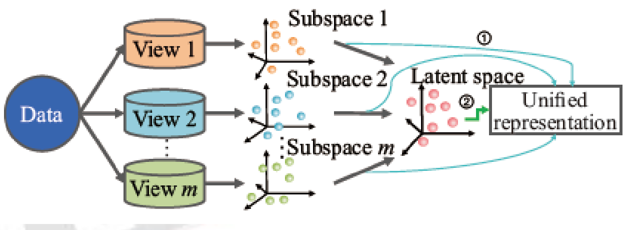

Figure 6: General procedure of multi-view subspace clustering.

The only thing that the latent space is is an illustration of the data after it has been compressed. During the compression process, the points of the data that are most similar to one another are brought closer together in space. Discovering less complicated ways to represent the data and gaining knowledge about the features of the data are both valuable uses of latent space. A unique method known as Latent Multiview Subspace Clustering (LMSC), which simultaneously analyzes underlying complimentary information from several viewpoints while clustering data points with latent representation [45].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

There are four primary categories of multiview subspace clustering methods. These include subspace learning-based methods, Non Negative Matrix Factorization (NMF)-based methods (which are considered to be a specific example of subspace learning), Canonical correlation analysis, and Deep multiview clustering methods.

4.1 Subspace learning based

Learning that takes place in subspace MVC makes the assumption that data points are drawn from the latent subspace in an effort to discover a latent space that can be derived from numerous low-dimensional subspaces. In recent years, there has been a significant focus in research on computer vision on subspace learning, which is the most important factor in dimensionality reduction. It includes traditional linear dimensionality reduction approaches, as well as manifold learning and other similar processes. Co-training is a method that Zhao et al. [44] developed in order to instruct multi-view subspace clustering. This framework merged the traditional K-means clustering method with the linear discriminant analysis in order to create a co-training scheme. This scheme made use of labels automatically learnt in one view in order to build discriminative subspaces in another view.

4.2 Non negative matrix factorization (NMF)

Clustering is a typical application for NMF. It requires one of the factorized matrices to be used as an indicator matrix, meaning that any element in the matrix that is not zero can be used to determine which information point belongs to which cluster. Therefore, a natural technique to do MVC is to insist that the indicator matrix for different views have the same or comparable structure. One of the most well-known properties of NMF is that it is able to learn a part-based representation, which is made possible by the nonnegative constraints. In many different applications, such as pattern recognition, information retrieval, and facial recognition, it provides both an intuitive and meaningful experience.

4.3 Canonical correlation analysis (CCA)

The purpose of canonical correlation analysis is to find and quantify the relationships that exist between two different sets of variables. The correlation among the independent and dependent variables is to be used as a guide when doing the clustering process.Given the labels for clustering, it is assumed that all perspectives are conditionally independent of one another. Chaudhuri et al. [46] introduced a method for learning multi-view subspaces that is based on classical correlation analysis. Using this method will allow you to acquire auxiliary results for Gaussian mixes as well as log concave distribution mixtures. The convex subspace representations learning method that Guo[47] presented for MVC was developed further. The primary objective is to identify a shared subspace representations across various perspectives, and after this has been accomplished, to apply conventional clustering algorithms to the shared representation. A co-training approach for multi-view subspace clustering was developed by Zhao and colleagues [48]. It integrated the traditional K-means algorithm with the linear discriminant analysis in the context of a co-training scheme that made use of labels that were automatically learnt in one view in order to produce prejudiced subspaces in another view. Deng et al. [49] proposed a feature weighting approach that was based on subspace learning. This method automatically and locally changed the feature weighting of each group based on how close the views were.It is possible to use canonical correlation with two different interpretations of the data.

4.4 Deep multiview clustering

Deep multi-view subdomain cluster based on intact space learning model (DMVSC-ISL) was a concept that was proposed by Zhao, Ding, and Y. Fu in [50]. This model incorporates both a subspace clustering module as well as a space learning module that is complete. The learning module for intact space makes use of auto-encoder networks to obtain the latent representation of each view by reducing redundant and noisy information, and then makes use of degradation networks to produce the intact space representation based on the various latent representations. It is more time-consuming, but the end effect is superior [51].

5. Multi task Multi Viewclustering

Multi-task clustering, which is a subfield of the field of multi-task learning [52], combines the performance of various tasks that are connected to one another and makes use of the connection that exists between these tasks in order to improve clustering performance for single-view data. The Multi-task Multi-view Clustering (MMVC) algorithm considers each individual view data with either one task or several tasks, as demonstrated in Figure 7. Inheriting the inherent characteristics of both the multiview clustering and the multitask clustering methods allows for this to be achieved. Finding a way to convey the intra-task (within-task) grouping on each view and finding a way to exploit the multitask and multiview relationship while concurrently conveying the inter-task (between-task) information to one another are the two primary issues that are associated with MMVC. Both of these challenges require finding a way to represent the intra-task (within-task) grouping on each view. Both of these challenges need to be overcome. Both of these challenges must be overcome before MMVC can be considered successful.

The purpose of multiview multitask clustering is to determine the connections between distinct but related activities in order

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

to improve the performance of each individual clustering activity [53], [54]. The performance of the first task can be improved with multi-task clustering algorithms by employing the information from the performance of other related tasks [55]. Many data from the actual world exhibit dual-heterogeneity. In other words, the data for many individual activities can be represented by various views, and the data for many distinct tasks can be shared across several views [54]. As an illustration, the clustering of songs sung in English and Hindi is an example of a multi-task clustering problem involving two tasks. The songs themselves are represented by two views: the lyrics and the audio features. Multi-task multiview clustering is the word that's used to describe this kind of issue.

This category takes into account each view as having either one task or numerous tasks that are related to one another, transfers the inter-task knowledge to one another, and leverages multi-task and multiview interactions in order to increase clustering performance.

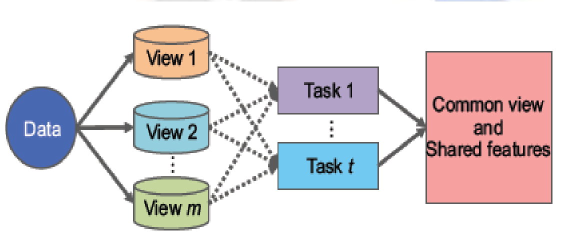

Figure 7: General procedure of multi-task multi-view clustering.

III. BENCHMARKS DATASETS

Seven benchmark multiview datasets are summarized in table I.

3Sources Dataset is a news article dataset. These articles are collected from three news sources: BBC, Reuters, and Guardians. There are 948 articles contained within the original datasets that are reported by at least one of the three different sources. 169 of these articles are included here, and a bag-of-words representation has been chosen to represent the articles as a whole. One of these six topical categories—business, entertainment, health, politics, sport, or technology—predominates throughout the majority of these 169 articles.

Reuters: The materials that are included in Reuters are translated into five different languages: English, French, German, Spanish, and Italian. Bag of words is employed to indicate the features that are present in each of the five perspectives that were created using these five language versions of these publications. These files fall under one of the six categories that are available. In order to generate a dataset consisting of 600 documents, a random sample of 100 documents is taken from each category.

Movies include 617 movies belonging to 17 genres. Each movie is described by two views: 1878 keywords and 1398 actors.

Yale is a widely used gray face dataset, which contains 15 categories and each category has 11 images. Three image features, i.e., the intensity feature (4096-D), the LBP feature (3304-D) and the Gabor feature (6750-D) are widely used as three types of views.

Caltech101 contains 8677 images of 101 categories. Each category has 40 to 100 images. There are three popular multi-view datasets generated from Caltech101, which are, respectively, Caltech101-7, Caltech101-20 and Caltech101. The Caltech101-7 and Caltech101-20 datasets with the size of 1474 and 2386 are generated from the subset of Caltech101, where 7 and 20 categories are, respectively, included. Both of them have six views consisting of the Gabor feature (48-D), the Wavelet-moment feature (40-D), the Centrist feature (254-D), the HOG feature (1984-D), the GIST feature (512-D), and the LBP feature (928-D). The Caltech101 dataset is composed of four views, whose dimensions are, respectively, 2048, 4800, 3540,1240.

Columbia Object Image Library (COIL20) contains 1440 images of 20 categories. The intensity feature (1024-D), the LBP feature (3304-D), and the Gabor feature (3304-D) were the three types of features that were utilized in order to provide an accurate description of the images. The Multiple feature handwriting digit dataset (Mfeat) stores the handwriting digital numbers from 0 to 9, with 200 handwriting digital photos for each number class. For the purpose of accurately representing an image, we choose to use the profle correlation feature (216-D), the Fourier coefcient feature (76-D), and the Karhunen–Loevecoefcient feature (64-D).

Table 1: Multiview Datasets

Dataset	Samples	Views	Features in eachview
3Sources	169	3	3560, 3631, 3068
Reuters	600	5	21526, 24892, 34121, 15487, 11539
Movies	617	2	1878,1398
YALE	165	3	4096, 3304, 6750
Caltech 101	8677	4	2048, 4800, 3540,1240.
COIL20	1440	3	30, 19, 30
Handwritten Digits	2000	3	76, 216, 64

Handwritten digit dataset is one of the benchmark datasets used for multiview clustering. Handwritten Digits is available from the UCI repository. This data set consists of following

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

three views of handwritten digit images, with classes 0-9. 76 Fourier coefficients of the character shapes

1. 216 profile correlations
2. 64 Karhunen-Love coefficients: Each image is encoded by its Karhunen-Love coefficients, a 64-dimensional vector. This results in a point cloud of 2000 points (2000 rows), living in 64 dimensions (64 columns)

IV. PERFORMANCE ANALYSIS

Normalized mutual information (NMI), accuracy (ACC), adjusted rand index (ARI), F-score, Precision, and Recall can be used to verify the performance of clustering. For all these metrics, the higher value indicates better clustering performance.

a) Normalized Mutual Information (NMI)

It is a normalization of the Mutual Information (MI) score to scale the results between 0 (no mutual information) and 1 (perfect correlation). It is used to determine the quality of clustering.

The formula of NMI is

$$NMI(Y,C) = \frac{2 \times I(Y;C)}{[H(Y) + H(C)]}$$

where,

1) Y = class labels

2) C = cluster labels

3) H(.) = Entropy

4) I(Y;C) = Mutual Information between Y and C

Note: All logs are base-2.

b) Rand Index

The Rand Index is yet another indicator that is frequently utilized. It does this by looking at every possible pair of samples and counting how many of those pairings are assigned to the same or different clusters based on the results of the predicted and the actual clustering. This gives it a measure of how similar the two clusters are to one another.

The formula of the Rand Index is:

R = (a+b) / (nC2)

where:

a: The number of times a pair of elements belongs to the same cluster across two clustering methods.

b: The number of times a pair of elements belong to difference clusters across two clustering methods.

${}_{n}C_{2}$**:** The number of unordered pairs in a set of *n* elements.

The Rand index always takes on a value between 0 and 1 where:

0: Indicates that two clustering methods do not agree on the clustering of any pair of elements.1: Indicates that two clustering methods perfectly agree on the clustering of every pair of elements.

c) Adjusted Rand Index

The Adjusted Rand Index (ARI) is a statistical measure that determines the degree to which two distinct data clusters are comparable to one another. It is a modification of the Rand Index, which is a fundamental method for determining the degree to which two clusterings are alike; nevertheless, it suffers from the drawback of being dependent on random occurrences. The Rand Index is a statistic that compares two different clusterings in terms of how similar they are to one another. The index is rescaled using the Adjusted Rand Index, which takes into account the fact that some objects will be able to occupy the same clusters simply because of random chance. As a direct consequence of this, the Rand Index will in fact never equal zero.

Performance of MVC techniques

The results are shown in table II

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

Table 2: Performance of five algorithms on six multiview datasets

Dataset	Method	Accuracy	F-score	Precision	Recall	NMI	ARI
3 sources (3 views)	Co-regularized multiview spectral clustering	0.54	0.46	0.49	0.44	0.49	0.32
	Multiview low rank subspace clustering	0.67	0.64	0.68	0.61	0.59	0.54
	Kernel –based weighted multiview clustering	0.36	0.36	0.23	0.84	0.11	-0.0064
	Multiview CCA	/	/	/	/	/	/
	Multiview clustering via deep matrix factorization	0.65	0.51	0.56	0.46	0.49	0.38
Reuters (5 views)	Co-regularized multiview spectral clustering	0.48	0.37	0.34	0.41	0.3	0.23
	Multiview low rank subspace clustering	0.52	0.42	0.36	0.49	0.38	0.28
	Kernel –based weighted multiview clustering	0.23	0.29	0.17	0.83	0.19	0.02
	Multiview CCA	/	/	/	/	/	/
	Multiview clustering via deep matrix factorization(Deep NMF)	0.29	0.22	0.21	0.23	0.11	0.06
Handwritten digits (3 views)	Co-regularized multiview spectral clustering	0.76	0.68	0.67	0.7	0.73	0.65
	Multiview low rank sparse subspace clustering	0.77	0.73	0.69	0.76	0.78	0.69
	Kernel –based weighted multiview clustering	0.87	0.75	0.74	0.77	0.77	0.73
	Multiview CCA	/	/	/	/	/	/
	Multiview clustering via deep matrix factorization(Deep NMF)	0.77	0.75	0.70	0.792	0.796	0.716
COIL 20 (3 views)	Co-regularized multiview spectral clustering	0.96	0.96	0.95	0.99	0.99	0.96
	Multiview low rank subspace clustering	0.98	0.98	0.97	0.99	0.99	0.98
	Kernel –based weighted multiview clustering	1	1	1	1	1	1
	Multiview CCA	/	/	/	/	/	/
	Multiview clustering via deep matrix factorization(Deep NMF)	0.39	0.27	0.21	0.37	0.51	0.22
YALE (3 views)	Co-regularized multiview spectral clustering	0.59	0.46	0.44	0.49	0.64	0.42
	Multiview low rank subspace clustering	0.58	0.41	0.39	0.44	0.61	0.37
	Kernel –based weighted multiview clustering	0.64	0.47	0.41	0.57	0.68	0.44
	Multiview CCA	/	/	/	/	/	/
	Multiview clustering via deep matrix factorization(Deep NMF)	0.74	0.56	0.55	0.58	0.73	0.54
Movies (2 views)	Co-regularized multiview spectral clustering	0.26	0.15	0.14	0.17	0.27	0.09
	Multiview low rank subspace clustering	0.32	0.19	0.19	0.20	0.32	0.14
	Kernel –based weighted multiview clustering	0.1	0.1	0.06	0.92	0.7	0
	Multiview CCA	0.13	0.06	0.06	0.06	0.085	0.0007
	Multiview clustering via deep matrix factorization(Deep NMF)	0.18	0.095	0.09	0.098	0.16	0.03

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
***Article Received**: 28 April 2023 **Revised:** 10 June 2023 **Accepted:** 27 June 2023*

"/" indicates this algorithm only applies to two-view case directly but this dataset has more than two views.

According to the findings in Table II, the deep MVC algorithm, the multi kernel MVC algorithm, and the multiview subspace clustering group algorithm all perform very well. On the six datasets that are most frequently used, the performance of Spectral clustering-based MVC, NMF-based MVC, and MVCCA is inferior to that of the algorithms described above.

V. CONCLUSION

The proliferation of multiview data necessitates the development of more complex clustering systems that can mine multiview datasets for hidden knowledge. The usage of multiview clustering has been demonstrated in a wide variety of real-world applications, including the detection of communities within social networks, the annotation of images within computer vision, cross-domain user modeling within recommendation systems, and the analysis of protein interactions within bioinformatics. The majority of the existing multiview clustering algorithms were analyzed and categorized in this study. These multiview clustering algorithms were placed into one of five broad categories, which are as follows: co-training style algorithm; multi-kernel learning; multiview graph clustering; multiview subspace clustering; and multi-task multiview clustering. The various approaches each come with their own set of benefits and drawbacks. To put it another way, algorithms that use a co-training approach can improve the clusters of different views dynamically by exchanging information with one another. On the other hand, they get challenging when there are more than three perspectives involved. The kernel-based multiview clustering approach makes use of the benefits offered by the kernel, but at the expense of increased computing complexity. The multiview graph clustering method employs spectral graph theory while relying on similarity matrices that have been built by the user. The methods of multiview subspace clustering are easy to grasp, but they also have a dependence on the initialization. Both the characteristics of multitask clustering and those of multiview clustering are inherited by multiview multitask clustering.

REFERENCES:

[1] P. Berkhin, "Survey of clustering data mining techniques," Yahoo, Sunnyvale, CA, USA, Tech. Rep., 2002, doi: 10.1007/3-540-28349-8_2.
[2] J. G. Saxe, The Blind Men and the Elephant. Hong Kong: Enrich Spot Limited, 2016.
[3] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," 2013, arXiv:1304.5634.
[4] S. Sun, "A survey of multi-view machine learning," Neural Comput. Appl., vol. 23, no. 7/8, pp. 2031–2038, 2014.
[5] J.Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recentprogress and new challenges," Inf. Fusion, vol. 38, pp. 43–54, 2017.
[6] S. Bickel and T. Scheffer, Multi-view clustering, in Proc. 4th IEEE Int. Conf. Data Mining, Brighton, UK, 2004, pp. 19–26.
[7] K. Nigam and R. Ghani, Analyzing the effectiveness and applicability of co-training, in Proc. 9th Int. Conf.Information and Knowledge Management, McLean, VA, USA, 2000, pp. 86–93.
[8] HamidrezaMirzaei "A Novel Multi-view Agglomerative Clustering Algorithm Based on Ensemble of Partitions on Different Views" 2010 International Conference on Pattern Recognition
[9] VikasSindhwani , ParthaNiyogi , Mikhail Belkin "A Co-Regularization Approach to Semi-supervised Learning with Multiple Views"Proceedings of the Workshop on Learning with Multiple Views, 22 nd ICML, Bonn, Germany, 2005
[10] Khetani, V. ., Gandhi, Y. ., Bhattacharya, S. ., Ajani, S. N. ., & Limkar, S. . (2023). Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), 253–262.
[11] Yang Li; Dapeng Tao; Weifeng Liu; Yanjiang Wang "Co-regularization for classification" Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics (SPAC) DOI: 10.1109/SPAC34359.201418-19 Oct. 2014
[12] Kumar, Abhishek, and Hal Daumé. "A co-training approach for multi-view spectral clustering."Proceedings of the 28th International Conference on Machine Learning (ICML-11). 2011.
[13] Nigam, Kamal, and RayidGhani. "Analyzing the effectiveness and applicability of co-training."Proceedings of the ninth international conference on Information and knowledge management.ACM, 2000.
[14] Zhou, Dengyong, and Christopher JC Burges. "Spectral clustering and transductive learning with multiple views."Proceedings of the 24th international conference on Machine learning.ACM, 2007.
[15] Sally El Hajjar, FadiDornaika,, "One-step multi-view spectral clustering with cluster label correlation graph," IEEE Trans. Knowl. Data Eng., vol. 21, no. 10,pp. 2022–2034, Oct. 2021
[16] L. Meng, A. H. Tan, and D. Xu, Semi-supervised heterogeneous fusion for multimedia data co-clustering, IEEE Trans. Knowl. Data Eng., vol. 26, no. 9, pp. 2293– 2306, 2014.
[17] J. W. Sun, J. Lu, T. Y. Xu, and J. B. Bi, Multi-view sparse co-clustering via proximal alternating linearized minimization, in Proc. 32nd Int. Conf. Machine Learning, Lille, France, 2015, pp. 757–766.
[18] G. Bisson and C. Grimal, An architecture to efficiently learn co-similarities from multi-view datasets, in Proc. 19th Int. Conf. Neural Information Proc., Doha, Qatar, 2012, pp. 184–193.
[19] G. Bisson and C. Grimal, Co-clustering of multi-view datasets: A parallelizable approach, in Proc. IEEE 12th Int. Conf. Data Mining, Brussels, Belgium, 2012, pp. 828– 833.
[20] S. F. Hussain and S. Bashir, Co-clustering of multi-view datasets, Knowl.Inf.Syst., vol. 47, no. 3, pp. 545–570, 2016.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 7s
DOI: https://doi.org/10.17762/ijritcc.v11i7s.7531
Article Received: 28 April 2023 Revised: 10 June 2023 Accepted: 27 June 2023

[21] Y. Jiang, J. Liu, Z. C. Li, and H. Q. Lu, Collaborative PLSA for multi-view clustering, in Proc. 21st Int. Conf. Pattern Recognition, Tsukuba, Japan, 2012, pp. 2997–3000.

[22] M. Ghassany, N. Grozavu, and Y. Bennani, Collaborative multi-view clustering, in Proc. 2013 Int. Joint Conf. Neural Networks, Dallas, TX, USA, 2013, pp. 1–8.

[23] V. R. De Sa, P. W. Gallagher, J. M. Lewis, and V. L. Malave, Multi-view kernel construction, Mach. Learn., vol. 79, nos. 1&2, pp. 47–71, 2010.

[24] V. R. De Sa, Learning classification with unlabeled data, in Advances in Neural Information Processing Systems 6, Cambridge, MA, USA, 1993, pp. 112–119.

[25] V. R. De Sa and D. H. Ballard, Category learning through multimodality sensing, Neural Comput., vol. 10, no. 5, pp. 1097–1117, 1998.

[26] S. Yu, L. C. Tranchevent, X. H. Liu, W. Glanzel, J. A. K. Suykens, B. De Moor, and Y. Moreau, Optimized data fusion for kernel k-means clustering, IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 5, pp. 1031–1039, 2012.

[27] M. Gonen and A. A. Margolin, Localized data fusion ¨ for kernel k-means clustering with application to cancer biology, in Proc. 28th Annu. Conf. Neural Information Proc. Systems, Montreal, Canada, 2014, pp. 1305–1313.

[28] Y. T. Lu, L. T. Wang, J. F. Lu, J. Y. Yang, and C. H. Shen, Multiple kernel clustering based on centered kernel alignment, Pattern Recognit., vol. 47, no. 11, pp. 3656– 3664, 2014.

[29] G. Tzortzis and A. Likas, Kernel-based weighted multiview clustering, in Proc. 12th Int. Conf. Data Mining, Brussels, Belgium, 2012, pp. 675–684.

[30] D. Y. Guo, J. Zhang, X. W. Liu, Y. Cui, and C. X. Zhao, Multiple kernel learning based multi-view spectral clustering, in Proc. 22nd Int. Conf. Pattern Recognition, Stockholm, Sweden, 2014, pp. 3774–3779.

[31] X. W. Liu, Y. Dou, J. P. Yin, L. Wang, and E. Zhu, Multiple kernel k-means clustering with matrixinduced regularization, in Proc. 30th AAAI Conf. Artificial Intelligence, Phoenix, AZ, USA, 2016, pp. 1888–1984.

[32] Y. Zhao, Y. Dou, X. W. Liu, and T. Li, A novel multiview clustering method via low-rank and matrix-induced regularization, Neurocomputing, vol. 216, pp. 342–350, 2016.

[33] P. R. Zhang, Y. Yang, B. Peng, and M. J. He, Multi-view clustering algorithm based on variable weight and MKL, in Proc. Int. Joint Conf. Rough Sets, Olsztyn, Poland, 2017, pp. 599–610.

[34] A. Trivedi, P. Rai, H. Daume III, and S. L. DuVall, ´ Multiview clustering with incomplete views, in Proc. Workshop on Machine Learning for Social Computing, Whistler, Canada, 2010. 35. W. X. Shao, X. X. Shi, and P. S. Yu, Clustering on multiple incomplete datasets via collective kernel learning, in Proc. 13th Int. Conf. Data Mining, Dallas, TX, USA, 2013, pp. 1181–1186.

[35] X. W. Liu, M. M. Li, L. Wang, Y. Dou, J. P. Yin, and E. Zhu, Multiple kernel k-means with incomplete kernels, in Proc. 31st AAAI Conf. Artificial Intelligence, San Francisco, CA, USA, 2017, pp. 2259–2265

[36] S. F. Hussain, M. Mushtaq, and Z. Halim, Multi-view document clustering via ensemble method, J .Intell.Inf.Syst., vol. 43, no. 1, pp. 81–99, 2014.

[37] J. C. Ni, H. H. Tong, W. Fan, and X. Zhang, Flexible and robust multi-network clustering, in Proc. 21th ACMSIGKDD Int. Conf. Knowledge Discovery and Data Mining, Sydney, Australia, 2015, pp. 835–844.

[38] V. R. De Sa, Spectral clustering with two views, in Proc. 22ndWorkshop on Learning with Multiple Views, Bonn, Germany, 2005, pp. 20–27.

[39] V. R. De Sa, Learning classification with unlabeled data,in Advances in Neural Information Processing Systems 6,Cambridge, MA, USA, 1993, pp. 112–119.

[40] V. R. De Sa and D. H. Ballard, Category learning throughmultimodality sensing, Neural Comput., vol. 10, no. 5, pp.1097–1117, 1998.

[41] D. Y. Zhou and C. J. C. Burges, Spectral clustering andtransductive learning with multiple views, in Proc. 24thInt. Conf. Machine Learning, Corvalis, OR, USA, 2007,pp. 1159–1166.

[42] Mohan, D. ., Ulagamuthalvi, V. ., Joseph, N. ., & Kulanthaivel, G. . (2023). Patient-Specific Brain Tumor Segmentation using Hybrid Ensemble Classifier to Extract Deep Features. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 127–135. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2579

[43] Y. Cheng and R. L. Zhao, Multiview spectral clusteringvia ensemble, in Proc. 2009 IEEE Int. Conf. GranularComputing, Nanchang, China, 2009, pp. 101–106.

[44] X. R. Zhao, N. Evans, and J. L. Dugelay, A subspace co-training framework for multi-view clustering, Pattern Recogn.Lett., vol. 41, pp. 73–82, 2014.

[45] C. Zhang et al., "Generalized latent multi-view subspace clustering,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 1, pp. 86–99, Jan.2020.

[46] K. Chaudhuri, S. M. Kakade, K. Livescu, and K. Sridharan,Multi-view clustering via canonical correlation analysis, inProc. 26th Annu. Int. Conf. Machine Learning, Montreal,Canada, 2009, pp. 129–136.

[47] Ekaterina Katya, S.R. Rahman. (2020). Void Node Detection and Packet Re-routing in Underwater Wireless Sensor Network. International Journal of New Practices in Management and Engineering, 9(04), 01 - 10. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/93

[48] Y. H. Guo, Convex subspace representation learning frommulti-view data, in Proc. 27th AAAI Conf. ArtificialIntelligence, Bellevue, WA, USA, 2013, pp. 387–393.

[49] X. R. Zhao, N. Evans, and J. L. Dugelay, A subspaceco-training framework for multi-view clustering, PatternRecogn.Lett., vol. 41, pp. 73–82, 2014.

[50] Sakura Nakamura, Machine Learning in Environmental Monitoring and Pollution Control , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[51] Q. Deng, Y. Yang, M. He, and H. Xing, Locally adaptivefeature weighting for multiview clustering, in Proc. 12thInt. FLINS Conf. Uncertainty Modelling in Knowledge Engineering and Decision Making, Roubaix, France, 2016,pp. 139–145.

[52] H. D. Zhao, Z. M. Ding, and Y. Fu, Multi-view clustering via deep matrix factorization, in Proc. 31st AAAI Conf. Artificial Intelligence, San Francisco, CA, USA, 2017, pp. 2921–2927

[53] Xu Yang, Cheng Deng, Zhiyuan Dang, and DachengTao,DeepMultiview Collaborative Clustering, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. 34, NO. 1, JANUARY 2023

[54] Q. Q. Gu and J. Zhou, Learning the shared subspace for multi-task clustering and transductive transfer classification, in Proc. 9 th IEEE Int. Conf. Data Mining, Miami, FL, USA, 2009, pp. 159–168.

[55] Z. H. Zhang and J. Zhou, Multi-task clustering via domain adaptation, Pattern Recogn., vol. 45, no. 1, pp. 465–473, 2012.

[56] S. N. Xie, H. T. Lu, and Y. C. He, Multi-task co-clustering via nonnegative matrix factorization, in Proc. 21st Int. Conf. Pattern Recognition, Tsukuba, Japan, 2012, pp. 2954–2958.

[57] S. Al-Stouhi and C. K. Reddy, Multi-task clustering using constrained symmetric non-negative matrix factorization, in Proc. 2014 SIAM Int. Conf. Data Mining, Philadelphia, PA, USA, 2014, pp. 785–793.

[58] A. Serra, D. Greco, and R. Tagliaferri, Impact of differentmetrics on multi-view clustering, in Proc. 2015 Int. JointConf. Neural Networks, Killarney, Ireland, 2015, pp. 1–8.

[59] Z. Xue, G. R. Li, S. H. Wang, C. J. Zhang, W. G. Zhang,and Q. M. Huang, GOMES: A group-aware multi-viewfusion approach towards real-world image clustering, inProc. 2015 IEEE Int. Conf. Multimedia and Expo, Turin,Italy, 2015, pp. 1–6.

[60] F. P. Nie, J. Li, and X. L. Li, Parameter-free auto-weightedmultiple graph learning: A framework for multiviewclustering and semi-supervised classification, in Proc. 25thInt. Joint Conf. Artificial Intelligence, New York, NY,USA, 2016, pp. 1881–1887.