Multiple Kernelized Fuzzy C-Means for Heterogeneous Data Using Unsupervised Machine Learning

Ankit R Mune[a], Dr.Sohel A Bhura[b]

[a]Assistant Professor of babasaheb naik college ofengineering, karla road, pusad, Maharashtra
[b]Assistant Professor of babasaheb naik college ofengineering, karla road, pusad, Maharashtra

Abstract: In the context of big data, heterogeneity is a prominent characteristic, & heterogeneousdata adds to difficulties of information convergence for big data analytics. The majority of datasets are diverse in terms of shape, formation, layout, variance, punctuation, and availability, among other characteristics. The broad variety of data sources usually results in software silos, which are a collection of unintegrated data systems with diverse methods, query languages, and application programming interfaces (APIs). During pre-processing stage, it is often necessary to combine data types from disparate sources. Scalability across diverse sources of information can only be achieved by automated or minimal human interaction in holistic models of data processing. Data aggregation is a method that brings together several different local resources without storing their information in a single central repository. Kernel learning is an active study area in the ML field, with many researchers workingon it at the same time. Over the last decade, researchers have conducted substantial research into the family of kernel-based ML techniques.ELM is a rapidly increasing-learning algorithm for a single hidden layer of feedforward neuralnetworks that may be used in both regressions as well as classification. ELMs for one hidden layer may randomly choose the node number of the hidden layer, as well as distribute input weights & hidden layer biases across the network. A mathematical modification is essential to finish the learning experience once the weights of the o/p layer have been established using the least square technique.
Keywords: Heterogeneous data, Multiple Kernel Learning, Extreme Learning Machine, Machine Learning.

1. Introduction

Heterogeneous data is definite as any data containing a varied area of information types as well as forms. This means they may be unclear and of poor quality owing to data redundancy, missing values, or outright lies. To satisfy corporate information needs, it is challenging to combine diverse data. For instance, Internet of Things (IoT) devices often creates heterogeneous data. The following four characteristics are often seen in IoT data. They have a wide variety of interests. The obtained data are also heterogeneous due to the wide range of data-gathering equipment. First & foremost, they're massive. It's important to keep track of all of the data that has been collected during a certain period, not just the most recent data. It's important to note that time & space have a close relationship. There is a time stamp on every item of data that comes in from data collection devices. A key feature of IoT data is a link between time and location. Fourth, effective data is just a tiny fraction of all the data that's out
there. Through the collection & transmission of data in IoT, a large number of sounds may be gathered. Datasets collected by acquisition devices only include a tiny fraction of the totalamount of

ISSN: 0010-8189

www.converter-magazine.info

information that may be useful. The forms of data heterogeneity listed below are a few examples **(Wang, 2017).**

Heterogeneous data is data that has a variety of kinds, structures, and distributions. It may bein a variety of forms, including Instagram, Facebook, Twitter, and YouTube, along with text, videos, photos, & some other sort of data, among other things. Diverse heterogeneous (mixed structured/ multi-structured) data sources, such as behavioural or social schemes, are often used in research. It may be composed of split values, heterogeneous characteristics, non-identical data sub-set distributions, or heterogeneous objects, among other things **(Cao, 2014).**

2. Challenges in Heterogeneous data

The structures and sizes of heterogeneous data are diverse. Managing all of these varied dataformats takes a lot of effort and money. Unstructured data, on the other hand, is far harder to process than structured data. Evaluating big amounts of diverse data is difficult since it takes a long time. More than one kind of data, such as email and other forms of communication, as well as data from sensors and other sources. Row or column storage may not be an option for this data. Transforming unstructured data into structured data is a difficult and &-consuming task. To efficiently manage this data, we'll need new technology **(Parimala*et al.*, 2017).** As a consequence, new techniques are needed. The primary reason for this is that learningtechniques may be better integrated when diverse knowledge is used. Nonsensical answers may be generated by a deep network without the application of a significant degree of supervision. Heterogeneous data training is time-consuming because of the enormous number of items involved. It is, nevertheless, an NP-hard problem to deal with heterogeneous information **(Last,2006).**

3.Kernel Learning (KL)

Kernel learning is an active study area in the ML field, with many researchers workingon it at the same time. Over the last decade, researchers have conducted substantial research into the family of kernel-based ML techniques (Breneman, 2005). KLs are a class of methodsfor pattern analysis in ML, with the support-vector machine (SVM) being a well-known member of the family. When dealing with datasets, the main impartial of pattern analysis is to discover and explore broad sorts of relations (for instance clusters or ranks, principal components, correlation or classification) that exist between them (or between them and each other). To address these problems, several more methods give that data points in raw portrayal are converted in feature vector depictions using a user-specified feature map; kernel techniques, on the other hand, do not require anything more than a user-specified kernel, which is a similarity function over pairs of information points in raw representation. SVM, Kernel LR, &Kernel PCA for denoising are just a few of the well-known instances of machine learning techniques. These kernel approaches have been successfully useful to a range of practical uses, with the efficiency of these techniques often showing promise. The kernel is the most important component of a kernel technique, and it is often defined as a function that provides an inner product at all between two samples in any induced Hilbert space, in particular. In general, by mapping data from input space to RKHS, which may be theoretically high-dimensional, standard linear algorithms are capable of being significantly improved in efficiency with a moderate investment of time and effort. Numerous practical investigations have shown that the choice of the kernel has a considerable impact on the ensuing performance of kernel techniques in many cases. Using the wrong kernels might indeed result in sub-optimal or very bad performance. Finding a suitable kernel function for many real-world scenarios is not always a straightforward undertaking; in fact, it may need significant domain expertise that would be problematic for non-expert users to get without extensive training. To overcome this constraint, researchers have been working hard in recent years to develop effective kernels that can be learned automatically from data (Bach, Lanckriet and Jordan, 2004)(Zhuang, Tsang and Hoi,2011b). Kernel learning may be used for a variety of tasks ranging from estimating the width parameters of a homogeneous Gaussian kernel to determining the optimum linear model of a collection of candidate kernels. The use of candidate sets of kernels or matrices offers an easymeans of integrating heterogeneous data into a single kernel machine, as has been proven in a no. of applications such as e.g., image processing(Girolami and Rogers, 2005). Known as kernel techniques in machine learning, they are a collection of methodologies that generate predictions about previously unseen data based on their resemblance to observations in the training set. It is possible to express pairwise similarity in terms of a function known as a kernel function. The prediction capacities of kernel techniques may be influenced by adjusting the kernel and its parameters, among other things. Kernel techniques are utilised in a broad range of applications, including classification, regression, or density estimation, amongst many others(Learning-machine, no date).

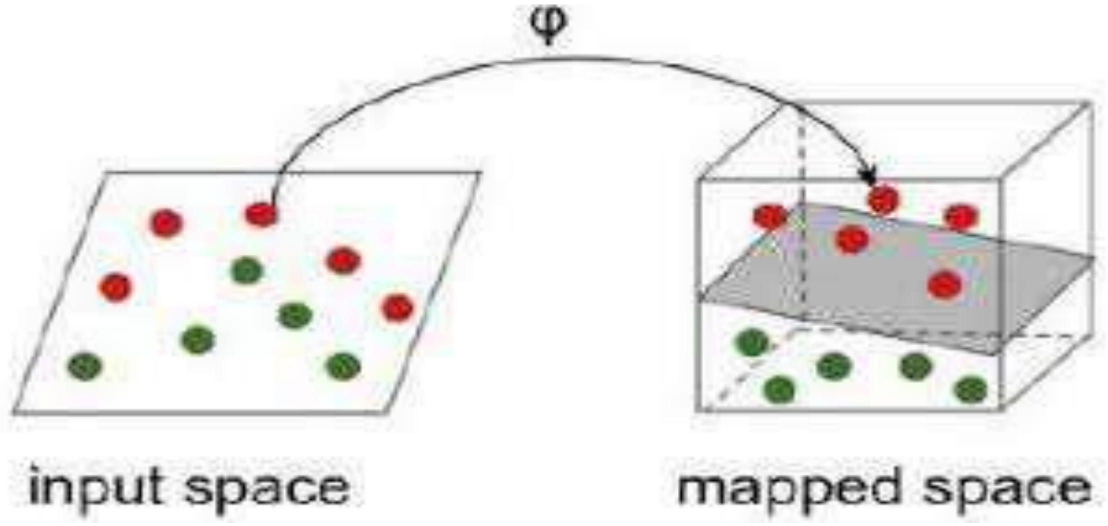

Figure.1: Kernel method

4. Multiple Kernel Learning (MKL)

In recent years, numerous approaches have been developed to associate several kernelsrather of using a single one, rather than a single one being used. It is possible that these different kernels correlate to the use of different conceptions of similarity, or that they are employing information derived from distinct sources (different representations or different feature subsets). When the optimal mix of characteristics is not known, learning linear groupings of many kernels is an intriguing method to use. For the task at hand, MKL is employed when there are diverse sources (representations) of data to be analyzed. MKL is an ML technique that seeks to build kernel machines for addressing real-world ML problems (for example, classification) by experimenting with different combinations of multiple kernels. The classic MKL technique is "shallow" in the sense that the target kernel is just linear (or convex) grouping of a no. of base kernels, which is a "shallow" approach. MKL is a prominent example of a kernel learning approach that seeks to learn linear (or convex) combinations of a collection of present kernels to select a desirable target kernel for application-specific applications. When compared to conventional kernel approaches that use a single fixed kernel, MKL demonstrates its superiority in terms of automated kernel parameter adjustment as well as its capacity to concatenate diverse data sources. The efficiency of MKL has been the subject of intense investigation over the last several years, with several methods being presented to resolve the

Performance of MKL, as well as several expanded MKL approaches being proposed to enhance the regular linear MKL method, among other things.

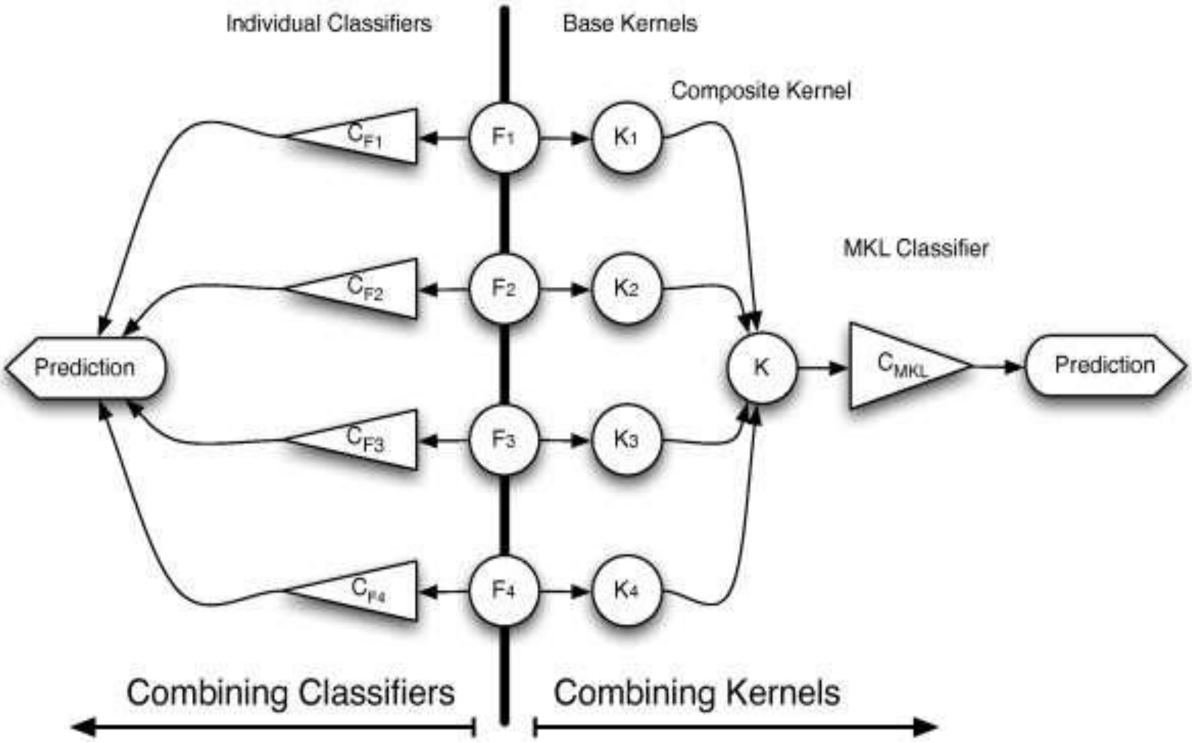

Figure 2: The intuition behind Multiple Kernel Learning

Multiple Kernel Learning is based on the same principles as Classifier Combination techniques,however, there are some variations. The Fs signify the various feature spaces which are being merged, and Ks indicate various kernels that are being compared with complete classifiers CFs, which are revealed on the left-hand side of the picture on the right-hand side of the figure (Polajnar, Damoulas and Girolami, 2011).

One of the most challenging aspects of any kernel-based algorithmic operations is aselection of an appropriate kernel & the calculation of parameters that define it. A no. of diverse kernels is tried on a validation set that is distinct from training data, but the one that performed the best is selected. To be clear, this is not an approach that can be utilized generally; it is time consuming, and it is certainly not appealing philosophically. Ideally, a collection of alternative kernels (including the possibility of the same kernel with various parameters) would be available, as well as the optimization method

ISSN: 0010-8189

could decide which kernel, or combination of kernels, is the best suitable. This is the scope of a research endeavour, which is often referred to as MKL (Theodoridis, 2020).

5. Unsupervised Multiple Kernel Learning (UMKL)

Is there a purpose to simply using unlabelled data for evaluation? UMKL is more difficult than traditional MKL problems since it does not incorporate class marks in the learning function. Using two intuitions to understand numerous kernels creates an uncontrollable situation:
• Each instance must be accurately recreated with its localized bases and kernel value weighted from a good kernel;
• Every instance must be adequately rebuilt with its localized bases and kernel valuation from a quality kernel...
UMK Function (Abbasnejad, 2011):

•Linear kernel: It's common to use a linear kernel in situations when the information may be split thru a single line. As the most widely utilized kernels, it's easy to see why. When there are a significant number of applications in a single data collection, this technique is often used. Each letter of the alphabet has its unique property when it comes to text categorization. There is a slew of options available. So, when it comes to text categorization, linear kernels are the most common.

•Polynomial Kernel: To do a dot calculation without converting the vector. Dot products x1 and x2 are computed in this manner as if they were a greater dimension of the original two vectors. A dot product from a separate function space is calculated differently by a kernel function.

•Gaussian Kernel: Regularity makes a Gaussian kernel an attractive choice for many density assessment applications. Since the approach does not enable numerous kernels to be mixed, all characteristics must be considered. Statistics tell us that a Gaussian distribution may be used as a standard for determining averages (proportionally high & heavy).

5.1 Challenges of UMKL

Due to the lack of class distinctions, a task of UMKL is more challenging than a regular MKLassignment. In contrast to MKL, UMKL's main difficulty is to find the optimal kernel k from fully unlabeled training data. To put it another way, the kernel's learning function must be guided by rules or intuitions that are independent of classes (Mariette and Villa-Vialaneix, 2018). Unsupervised MKL algorithms take a long time to improve. The numerical explanation utilized in the optimization of kernel combination coefficients is the primary reason for a slow learning rate. Management of large amounts of data and real-time learning is not met by this system.

6. Extreme Learning Machine (ELM)

ELM is a rapidly increasing-learning algorithm for a single hidden layer of feed forward neuralnetworks that may be used in both regressions as well as classification. It is implemented in C++. In feedforward network training, ELMs for one hidden layer may randomly choose the node number of the hidden layer, as well as distribute input weights & hidden layer biases across the network. A mathematical modification is essential to finish the learning experience once the weights of the o/p layer have been established using the least square technique. In comparison to the conventional BP method, the training time has been significantly improved. It first trains the ELM and afterward uses it to predict real-world applications. The data set used for training is generally comprised of issues that are unique to the data set. A variety of realworldoutcomes as well as related factors are included in the dataset. Impact factors and outcomes will be entered into ELM for the training stage using an iterative methodology to ensure that iteration is completed to finish the learning process. Then, only the input data set and the training data set are the same as the influencing variables for the ELM that has been trained. The ELM model will be determined by the results of the predictions made by memory. ELM for a single hidden layer feedforward neural network is a powerful and simple-toimplementneural network architecture. Although the old conventional NN learning algorithm (for example, the BP algorithm) necessitated the definition of multiple fake n/w training parameters, it may also lead to a locally optimal solution in a short time (Albadr and Tiun, 2017).

A new technique called ELM has been developed for training SLFNs. When using ELM,random initiation of hidden nodes is followed by fixation of hidden nodes without any need for repetitive tweaking. Furthermore, neurons do not have to be buried nodes to function properly in ELM. The only free variables to be known are the connection (or weights) between the hidden and output layers layer. As a result, ELM is built as a linear model which lowers it to a linear system resolution. ELM is incredibly successful, and it aspires to reach the best overall performance when compared to typical FNN learning methods. A theoretical investigation has demonstrated that ELM preserves the universal approximation capability of SLFNs even when the hidden nodes are freely generated by the user. ELM can get a nearly perfect generalization bound of a typical FNN with routinely utilized activation features in which most of itsparameters are known by using ELM's generalization bounding technique. The benefits of ELM over standard FNN algorithms in terms of efficiency and generalization effectiveness have been proven with a wide range of issues from a wide variety of different domains. It is vital to highlight that ELM is far more successful than SVM in most cases, while least-square SVM is one of the most up-to-date techniques available. Experimental evidence has indicated that ELM is equivalent to, or even superior to, SVMs and their derivatives when it comes to generalization ability.

6.1. Problem in ELM

• ELM is a layer of FFNNs that is not visible. Semi-controlled ELM (SELM) has beenincluded to solve the issue of unlabelled data. In this case, labelled data is also scarce or costly.

• Real-world applications, such as medical diagnostics & CCFD, are hampered by unbalanced data.

• Pattern identification will suffer if the ELM classification network has too few or too numerous hidden nodes.

7. Machine Learning Tasks and Algorithms

A variety of ML methods, e.g. classification and regression analysis, data clustering andlearning by association rules, feature engineering for dimensionality reduction, and DL approaches are discussed in this section. According to Fig. 1.14, a machine learning-basedpredictive model is separated into 2

ISSN: 0010-8189

stages: phase 1 involves training of model using historical data and phase 2 involves generation of results using fresh test data. (Sarker, 2021).

• **Naive Bayes (NB):** This approach is built on Bayes' theorem & assumption ofindependence among every pair of characteristics, and it is used to detect false positives. Numerous real-world applications, e.g. document or text categorization, spam filtering, and so on, benefit from its effectiveness and versatility. It may be applied for both binary & multi-class categories in these circumstances. Use of NB classifier is recommended for efficiently classifying noisy instances in data as well as for construction of a strong prediction model. The key benefit of this strategy is that, in contrast to more advanced systems, it requirements a modest quantity of training data to estimation essential parameters and can do so fast and accurately. Nevertheless, because of its high assumptions on feature independent, its efficiency may be adversely affected. “Gaussian, Multinomial, Complement, Bernoulli, as well as Categorical classifiers” are most often seen versions of the NB classifier.

• **K-nearest neighbours (KNN**): KNN has been widely used as a classification algorithm. NN is a classification strategy that categorises unknown data points by comparing them to surrounding data points whose values have previously been established. In the case of photo libraries and online marketing, for example, it may be used to find trends, as well as for cluster analysis. Because the K-NN approach is straightforward to construct, it is simple to integrate and debug activities. The user may choose to inspect the gate that follows. The most significant benefit of this strategy is that training may be completed more quickly, easily, and conveniently than with other methods. Reliable approaches may be used to determine large amounts of training data. It is essential to have vast data sets to educate. As with linear and nonlinear classifications, the k-nearest the neighbours' technique takes an original and innovative approach. The majority of algorithms, such as linear and nonlinear classifiers, make use of training data to adjust their internal structural elements. Classifiers based on K-nearest neighbours store training data and use it to categorise new characteristics based on the information in the training data. As a consequence, this classifier is often referred to as being very user-friendly by its users.

• **Support vector machine (SVM):** SVM is other prominent ML approach that may be utilized for classification, regression, & other tasks in addition to classification and regression. SVM is computer programme that generates a hyper-plane. Naturally, the hyper-plane that is the farthest away from any given training data point in just about any class gets important parting since larger the margin of error between two classifiers means that classifier's generalisation error is smaller. It is useful for high spaces and may act in a diversity of conducts depending on mathematical functions used, which are referred to as the kernel. Linguistic kernel functions, e.g. linear, polynomial, RBF, sigmoid, and others, are often utilised in SVM classifiers. SVM, on the other hand, does not function well if data set includes additional noise, including such overlapping target classes.

8. Data Clustering Approaches

Data clustering is among most important, frequent, and intriguing jobs in thecategorization of patterns in many fields including such “data mining, pattern recognition, artificial intelligence”, and so forth. The goal of data clustering is to categorise items that aresimilar to one another. It is possible to use a wide variety of various data clustering approaches, each tailored to a certain kind of application. Data clustering is a method in which we group items together based on their similarities in terms of characteristics. Data analysis is a prevalent tool in current scientific research, and it can be found in fields as diverse as communication and media studies, computer science, and biological science. Clustering, being the fundamental component of data analysis, plays a key part in the process. (Xu and Tian, 2015). On the one hand, a multitude of techniques for cluster analysis have been developed in tandem with the expansion in information and the intersection of subjects. Using a decent clustering approach, you may generate high-quality clusters that are highly similar inside their respective classes and have little inter-class similarity(Bano and Khan, 2018).

•**Hierarchical Clustering Method:** It is possible to generate a dendrogram, which is adiagram-like tree which preserves patterns of joins & separations, by using this approach. A dendrogram is a set of nested clusters that is organised as a visualised &hierarchical tree. Hierarchical methods may be agglomerative or divisive in whatever way they like. Every element is started as a single cluster, and then they are joined together in succession to form bigger clusters; divisive techniques begin with full group and continue to divide it in increasingly minor groups.

• **Partitioning Clustering Method**: Most of time, partitioning Procedures outcome in a collection of M clusters, with each item going to a distinct cluster as a consequence. For each cluster, a centroid or a cluster representative may be used to indicate it; this serves as a form of summary explanation of all the items included inside a cluster. According on the type of the object that is being clustered, the precise format of this report will vary.

• **Density-based Clustering Method:** This technique is capable of discovering clusters of any size and form. Objects are divided into dense and low-density regions based on their density. Connecting points are established on the basis of a predetermined distance criterion. Points are joined together until they meet the density criterion. It can deal with noise and just one scan. DBSCAN, OPTICS, DENCLUE, CLIQUE, BIRCH, and CURE are some of the density-based techniques that are available

•**Fuzzy Clustering:** In the field of soft computing, fuzzy clustering is used, as well asthe data points might be found in more than one cluster. The fuzzy c-means clusters technique is a famous fuzzy clustering method that is often employed in image processing applications. FCM is an iterative approach that is used to find cluster centres that have the lowest dissimilarity function. Because FCM employs fuzzy partitioning, a given piece of information may be assigned to a variety of groups. A squared error may be required to get a high-quality categorization. The prescription of fuzzy portioning is accomplished by iterative optimization of the goal. In both the k means or FCM clustering algorithms, the clustering mechanism is the same, however when the k means method clusters, it uses the mean of the weighted cluster, which makes it easier to locate masses or the

ISSN: 0010-8189

starting place of cancer or growth. In FCM, it is taken into account that each point has a weighted value that is tied to the cluster. This technique assisted physicians or radio logisticians in determining the fraction of breast cancer cases that had been opened up. The performance is dependent on the number of cluster centres that are first created. Outliers or noise are also present in the FCM data, making it difficult to distinguish between the first partitions and the subsequent partitions. The FCM method produces better results than the difficult k-means technique. (Of and In, 2020).

Also, the FCM clustering method has been applied in the field of picture segmentation. The fuzzy cmeans technique employs iterative optimisation of an objective function defined as the weighted measure of similarity here among pixels in the picture with each of the c-cluster centres in order to get the best possible result. A local extreme of this objective function impliesthat the input data has been clustered in the most optimum way. According to this formula, the objective function that is minimised is (Nagalakshmi and Jyothi, 2013):

$$W_m(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{C} (\mu_{ik})^m \; (d_{ik})^2 \; .$$

The "μik fuzzy membership value of the Kth pixel in ith cluster, d ik is really any inner product-induced norm metric, m controls essence of clustering, with hard clustering at m=1 aswell as growingly fuzzier clustering at greater values of m, V is the set of c-cluster centres", as well as U is fuzzy c-partition of picture. In this equation: Using fuzzy set theoretic picturessegmentation, Trivedi &Bezdek developed a method for aerial image segmentation. The method is built on perceptions of region expanding and makes use of a pyramid data structure. The algorithm is structured in a hierarchical manner.

9. Multiple Kernelized Fuzzy C-means

Kernel approaches are a class of algorithms for pattern analysis in machine learningthat is used to identify the overall types of relationships in datasets. Kernel techniques to find the overall types of relationships in datasets. Many of the techniques used to solve these problems require that the raw-presented datasets be explicitly converted into the feature vector — presented ones — through a user-specified feature map before they can be used. When using kernel techniques, the fundamental concept is to invert the sequence of parameters, i,e, choose the appropriate Kernel K (rather than a mapping) before performing a learning process. There are many kernels. The fuzzy c-means algorithm offers us a novel technique for combining diverse information from unsupervised heterogeneous data. The information included in heterogeneous data is merged in the kernel space by mixing distinct kernel functions specified on different information domains. FCM clustering is a soft clustering strategy in that each data point is allocated a likelihood or probability score to indicate whether or not it belongs to a certain group. The Kernelized Fuzzy C-Means algorithm (KFCM) is established by altering the Vectors standards, Fuzzy Clustering Center, and partitioning matrices from Bezdek's conventional FCM in 1981. It defines as follows:

Step 1: Select the cluster number, c, $2 \leq c < n$, where n is collecting total data points or the number of feature vectors in the given dataset. Select m, $1 \leq m < \propto$. Set the standard vectors ‖ ‖ (here Multiple Kernel Function (including linear, polynomial, Gaussian) instead of Euclidean distance), i.e. (e.g., of Gaussian function)

ISSN: 0010-8189

$$k(x_r, v_i) = \sum_{t=1}^{p} e^{\left(\frac{-||x_{rt}, v_{it}||^2}{\sigma^2}\right)} \qquad (1)$$

Here,

x_{rt} = tth features of the rth feature vector, for r=1,2,3,4,..,n; t=1,2,3,4,..,p

v_{it} = t-dimensional center of the ith cluster for i=1,2,3,4,……,c; t=1,2,3,4,….,p;

n= total feature vectors,

p = no. of features in each feature vectors

c = total no. of clusters

Select the first fuzzy partitions (via putt few arbitrary values)

$$U^{(0)} = [\mu_{s_i}^{(0)}(x_r)]_{1\le i\le c, 1\le r\le n} \qquad (2)$$

Select a constraint >0 (this will tell us when to stop the iteration). Set up the parameter of iteration count l = 0.

Step 2: In this step of MKFCM, the fuzzy cluster centroids or centers$\{v_i^{(l)}\}_{i=1,2,3,4,......,c}$are computed by given eq. (3):

$$v_i^{(l)} = \frac{\sum_{r=1}^{n}\left(\mu_{s_i}^{(l)}(x_r)\right)^m k(x_r, v_i) x_r}{\sum_{r=1}^{n}\left(\mu_{s_i}^{(l)}(x_r)\right)^m k(x_r, v_i)} \qquad (3)$$

r= 1, 2,3,4,…..n.

Step 3: Compute the new partitioning matrices (that is, membership matrix) by using equation

$$U^{(l+1)} = [\mu_{s_i}^{(l+1)}(x_r)]_{1\le i\le c, 1\le r\le n} \qquad (4)$$

Here,

$$\mu_{s_i}^{(l+1)}(x_k) = \frac{\left(1-k(x_r, v_i^{(l)})\right)^{\frac{-1}{(m-1)}}}{\sum_{j=1}^{c}\left(1-k(x_r, v_i^{(l)})\right)^{\frac{-1}{(m-1)}}} \qquad (5)$$

for i=1,2,3,4,……..,c & r=1,2,3,4,……..,n.

If $x_r = v_i^{(l)}$ eq. (5) cannot be utilized. In this situation, the membership function is determined as:

$$\mu_{s_i}^{(l+1)}(x_r) = \begin{cases} 1 & if r = i \\ 0 & if r \neq i, \text{i} = 1,2, \dots\dots.., \text{c} \end{cases} \qquad (6)$$

Step 4: Calculate

$$\Delta = \left\|U^{(l+1)} - U^{(l)}\right\| \qquad (7)$$

If , then Continuesteps 2, 3 & 4. Otherwise, halt some count of iteration l^*.

10.Conclusion

Multiple sources without any information about labels gather a vast quantity of data,which is often heterogeneous, i.e. including a variety of kinds, structures, and distributions of data. Instagram, Twitter, Facebook, as well as YouTube, as well as texts, photographs, and videos, are examples of this kind of data. To obtain data from such large amounts of unlabelled heterogeneous data, advanced unsupervised learning algorithms (with many kernels) are used. MKL When it comes to revealing information from many sources, traditional MKL algorithms are a solid choice. Some attempts are now being undertaken to effectively collect data from heterogeneous distributed data that is difficult to manage, such as employing the kernel to manage difficult heterogeneous

ISSN: 0010-8189

distributed data. In recent years, several unsupervised learning methods have been conducted to improve learning outcomes.

References (APA)

1. Gu, Y., Feng, K. and Wang, H. (2017) ‘A new sparse multiple-kernel learning method for classification of hyperspectral imagery’, in Workshop on Hyperspectral Image and Signal Processing, Evolution inRemote Sensing. doi: 10.1109/WHISPERS.2013.8080695.
2. Cao, L. (2014) ‘Non-iidness learning in behavioral and social data’, Computer Journal. doi: 10.1093/comjnl/bxt084.
3. Parimala, K. et al. (2017) ‘Challenges and Opportunities with Big Data’, International Journal ofScientific Research in Computer Science and Engineering. doi: 10.26438/ijsrcse/v5i5.1620.
4. Last, M. (2006) ‘Kernel Methods for Pattern Analysis’, Journal of the American Statistical Association.doi: 10.1198/jasa.2006.s153.
5. Breneman, J. (2005) ‘Kernel Methods for Pattern Analysis’, Technometrics. doi: 10.1198/tech.2005.s264.
6. Bach, F. R., Lanckriet, G. R. G. and Jordan, M. I. (2004) ‘Multiple kernel learning, conic duality, andthe SMO algorithm’, in Proceedings, Twenty-First International Conference on Machine Learning,
7. ICML 2004. doi: 10.1145/1015330.1015424.
8. Zhuang, J., Tsang, I. W. and Hoi, S. C. H. (2011a) ‘A family of simple non-parametric Kernel learning algorithms’, Journal of Machine Learning Research.
9. Girolami, M. and Rogers, S. (2005) ‘Hierarchic bayesian models for kernel learning’, in ICML 2005 - Proceedings of the 22nd International Conference on Machine Learning. doi: 10.1145/1102351.1102382.
10. Polajnar, T., Damoulas, T. and Girolami, M. (2011) ‘Protein interaction sentence detection using multiple semantic kernels’, Journal of Biomedical Semantics. doi: 10.1186/2041-1480-2-1.
11. Theodoridis, S. (2020) ‘Learning in Reproducing Kernel Hilbert Spaces’, in Machine Learning. doi: 10.1016/b978-0-12-818803-3.00022-2.
12. Abbasnejad, M. E. (2011) ‘An unsupervised approach to learn the kernel functions: from globalinfluence to local similarity’, Neural Computing and Applications, 20(5), pp. 703–715. Available at: https://dl.acm.org/doi/10.1007/s00521-010-0411-7.
13. Mariette, J. and Villa-Vialaneix, N. (2018) ‘Unsupervised multiple kernel learning for heterogeneousdata integration’, Bioinformatics. doi: 10.1093/bioinformatics/btx682.
14. Albadr, M. A. A. and Tiun, S. (2017) ‘Extreme learning machine: A review’, International Journal ofApplied Engineering Research.
15. Sarker, I. H. (2021) ‘Machine Learning: Algorithms, Real-World Applications and Research Directions’, SN Computer Science, 2(3), pp. 1–21. doi: 10.1007/s42979-021-00592-x.
16. Dr.AarushiKataria, Dr. Naveen Nandal and Dr. Ritika Malik, Shahnaz Husain -A Successful Indian Woman Entrepreneur, International Journal of Disaster Recovery and Business ContinuityVol.11, No. 2, (2020), pp. 88–93
17. Kumar, S. (2020). Relevance of Buddhist Philosophy in Modern Management Theory. Psychology and Education, Vol. 58, no.2, pp. 2104–2111.

18. Roy, V., Shukla, P. K., Gupta, A. K., Goel, V., Shukla, P. K., & Shukla, S. (2021). Taxonomy on EEG Artifacts Removal Methods, Issues, and Healthcare Applications. Journal of Organizational and End User Computing (JOEUC), 33(1), 19-46. http://doi.org/10.4018/JOEUC.2021010102
19. Shukla Prashant Kumar, Sandhu Jasminder Kaur, Ahirwar Anamika, Ghai Deepika, MaheshwaryPriti, Shukla Piyush Kumar (2021). Multiobjective Genetic Algorithm and Convolutional Neural Network Based COVID-19 Identification in Chest X-Ray Images, Mathematical Problems in Engineering, vol. 2021, Article ID 7804540, 9 pages. https://doi.org/10.1155/2021/7804540
20. Aarushi, Naveen Nandal, Parul Agrawal. AN EXPLORATORY RESEARCH IN PRODUCT INNOVATION IN AUTOMOBILE SECTOR. JCR. 2020; 7(2): 522-529. doi:10.31838/jcr.07.02.98
21. Bano, S. and Khan, M. N. A. (2018) 'A Survey of Data Clustering Methods', International Journal of Advanced Science and Technology. doi: 10.14257/ijast.2018.113.14.
22. Of, O. and In, E. (2020) 'A SURVEY PAPER ON IMPROVING FUZZY C-MEANS', VIII(Iii).
23. Nagalakshmi, G. and Jyothi, S. (2013) 'A Survey on Pattern Recognition using Fuzzy Clustering Approaches', International Refereed Journal of Engineering and Science (IRJES) ISSN (Online).