ATBCWCE: Automated Two-level Variable Big Bang Big Crunch k Weighting Clustering Ensemble Framework for Multi view Temporal Data and Cancer Data

Kalaiarasu M.*and Anitha J.
Department of IT, Sri Ramakrishna Engineering College, Coimbatore, INDIA
*kalai.muthusamy@srec.ac.in

Abstract
Temporal data unsurprisingly take place in different kind up and coming applications like sensor networks, person mobility or internet of obsession. Result and experience dataset a sample is collected in huge databases are dependent upon temporal contacts among patients and their physicians and pharmacists. In this research work, the hypothetical effect of Temporal Clustering (TC) in various domains are examined which in turn will be useful for physician and pharmacy services on pharmaceutical risk estimates in cancer genome by introducing data mining clustering algorithms. For the first time, cancer genome datasets samples were made available for the complete genome sequences consisting of point mutations and structural alternations for a huge number of cancer types which allows the variation of cancer subtypes in an exceptional global analysis. In this work, TC algorithm is presented to the allocation of numerous time series based set of non-overlapping parts that fit with k temporal clusters. Specifically, present experimental studies expose that temporal data clustering having a major difficult in temporal data mining based on elevated dimensionality, weight value calculation, multi view data and multifaceted temporal correlation. Therefore, group of clustering gives a common enabling method to make use of dissimilar illustrations together for temporal data clustering. In the multi view temporal data these king of partitioning methods are not appropriate. For that reason, in the multi view data, the TW-k-means and an automated two-level variable clustering algorithms are used. At the same time it can be used to calculate the views weights and person variables. A new ATBCWCE structure is used to solve all the trouble of above mentioned methods for clustering multi view temporal datasets samples.

In the initial clustering analysis the ATBCWCE method is used to produce a multiple partitions and group of clustering structure and also it can be used to produce a final partition by grouping those partitions. After this process, the ATBCWCE structure, all the input partitions are grouped together. In the ATBCWCE structure, each and every view the view weight value is assigned to recognize the compression view and also the weight value has been allocated to the every variable. Here the distance is calculated by using view weights and variable weights based on which clustering would be done. At this time, the Big Bang Big Crunch algorithms are developed with extra two steps is included in iterative k-means clustering method to mechanically calculate the view based weight values and variable. The future ATBCWCE structure has been estimated with time series benchmark standard. In the temporal data clustering, the simulation results show that better result of ATBCWCE structure.

Index Terms: Clustering Ensemble, Temporal Data Clustering, Weighted Consensus Function, Different Representations, Multi view Learning, k-means, variable weighting, Big Bang Big Crunch.

Introduction

In all the pattern analysis tasks, clustering is the very significant responsibilities for the summarization purpose, clustering and the prototype mining, and also for the k-means++, k-means clustering fuzzy c-means, big data dimensionality reduction, and the other kind of variations in the middle of the very famous clustering algorithms, for the reason that it provides a better trade-off between the quality of the elucidation and also the difficulty of computational process [1].

In contrast, the Dynamic Time Warping (DTW) usually using a k-means clustering of temporal data otherwise a number of temporal kernels [2-3] is difficult for the reason that the clusters require a multiple temporal data concurrently. Some of the methods are limited in the standard DTW and some of the temporal data having the similar inclusive characteristics [4]. In temporal data clustering method to avoid the centroid estimation difficulties, kernel k-means [5] and k-means are commonly used [6]. And these methods are unsuitable for the capturing local temporal features [7].

Specifically, the present experimental studies [2-5] expose that the temporal data clustering having major difficulties in temporal data mining owing to multifaceted temporal correlation and high dimensionality. The temporal-proximity model and representation supported clustering algorithms are cast-off to the data dependency management and also to classify the previous temporal data clustering algorithms.

The two-step algorithms for clustering method were developed by Liao [8] for multivariate time series of equivalent or else uneven length. Initially k-means and FCM clustering algorithms are processed to find the real valued time series. This univariate discrete valued time series variable is understood as state variable method. The second phaseis utilized to form the transition probability matrices by grouping the univariate discrete valued time series in k-means or FCM algorithm into a numerous clusters.

The Normalized Longest Common Subsequence (NLCS) are used to analyse the time series. This method was implemented by Dacheng [9]. This NLCS method is usually used in comparing the character sequences. In this work, Dacheng was implemented a new NLCS algorithm to estimate the comparison of time series. This algorithm is commonly used for all the unique subsequence as an alternative for a highest general sequences correctly.

The data stream clustering methods was implemented by Wu [10] and it is used to analysis the stock data. For the period of clustering process the data stream clustering method intended to keep shape and have a propensity features. There are two segments are available in this clustering process. The first one is online clustering and the second one is offline macro clustering. Online clustering is used to maintain the clustering feature vectors and also it is used to extract the characteristics of data flow. Offline macro clustering process is used to act in response the user requirements. The better results are gained after completing this clustering process. The proximity method is applied to the temporal data and the proximity method performance took a more time to complete a task. For that reason only the clustering methods are used to the some of the work

The predicting agriculture deficiency process is analysed by Huang [11] supported fuzzy set and R/S investigational replica. Huang using a fuzzy clustering iteration technique and it is used cluster the data and also estimating the weight value. This process is affected the crop output in every enlargement phase. The produced outcome displayed the model is suitable and possible in the application in the agricultural deficiency.

The prediction accurateness has been increasing with adjusting the two deficiencies by Lin [12]. The sub intervals are used to represent the data distribution structures. First of all, Lin partitioning the sub intervals with midpoints of two nearest clusters centres and sub-intervals works on the fuzzy the time series. After that the fuzzy time series replica with multi factors fuzzy association is built-with stock data.

Autoregressive Conditional Heteroscedasticity (ARCH) [13] possessions of storm series of data are analysed with E views software. Initially, the storm data series are produced by using the Autoregressive Moving Average (ARMA) model. After the initial step, by using the Lagrange multiplier the ARMA replica is examined and equivalent ARMA-ARCH replica is set up. Finally, ARMA model process is compared to the ARMA-ARCH performance. Justification of ARMA-ARCH model is verified. A representation based algorithm used to convert the temporal data clustering into the static data clustering and this process is used to capture the data dependency.

The wave cluster using the financial time series for mining purpose and it was developed by Jixue [14]. This wave filter is one type of grid filter and also the density filter. Every theoretical and practical possibility of time series data mining depends on space renovation phase. Subsequent to time series pattern mining method, the Time Series Data Mining structure (TSDM) is noted down, and also the temporal patterns mining technique supported wave cluster is methodically obtainable.

Powell et al [15] will evaluate the unverified categorization techniques like k-means clustering with supervised learning algorithms like Support Vector Machines (SVMs). In this work, from a historical data of the S&P 500 a number of stock prices are engaged and it is used as test bed. These stock prices will be characterized as rising or diminishing in price on a periodical origin.

The grouping of Independent Component Analysis (ICA) with Support Vector Regression (SVR) was implemented by Wu et al [16] and it is used to predict the time series financially. In this work, the new Structure Risk Minimization (SRM) technique was used. It eliminates the difficulties of the learning algorithms based on Empirical Risk Minimization (ERM) method. It is only suitable for the training samples not for the future samples. In this case, feature extraction is done by using the ICA after that for the prediction process the non-linear SVR method is applied to complete the process.

Verdoolaege [17] was implemented a novel technique for detection of activated voxels in the BOLDFMRI data. Initially voxel time series is extracted from the wavelets using Generalized Gaussian Distribution (GGD). At last, in the GGD method the k-means clustering process is performed to the voxel information in the synthetic information set with the Kullback- Liebler Divergence (KLD) as a correspondence compute.

To characterize the each and every variety of temporal data there is no particular representation method. Additionally, it is very complicate to select the temporal data set appropriately without past information and watchful analysis process. These troubles frequently hold back a representation-based approach from the satisfactory presentation. For the multi view temporal dataset samples these methods are not suitable.

The upcoming region in the machine learning application, the clustering ensemble algorithms are presently studied in the various kind of view point. For example, evidence aggregation [19], graph partitioning [18] and semi definite programming optimization [20]. The essential thought of the following clustering ensemble algorithm is grouping of multiple partitions on the similar data set to generate consent particles. This consent particle is usual to be the better to that of specified input partitions. Even though, the theoretical justification on clustering ensemble algorithms, increasing experiential evidences maintain like thought and mention that clustering ensemble is proficient for detecting new cluster structures [18-20]. Additionally, in the clustering ensemble the official analysis exposes beneath some particular conditions, an appropriate consent result uncovers the inherent structure of specified information set [21]. Consequently, the clustering ensemble approach gives common methods to use a various representation together for the temporal data clustering.

From the various feature spaces, the multi view data is the variable that having the multiple views such as representation or group of variable. The result of combination of several kinds of measurements on examination from numerous perspectives and various measurement types can be referred as numerous views. In the precedent decade, the multi view information has increasing the security and it is called as multi view clustering [22].In the clustering method, from the set of variables it takes the numerous views and also eliminates the differences between different views.

From the multiple views, clustering reveals the data and gets the dissimilar between the various views into deliberations in order to generate an exact outcome and vigorous data partitioning. In cluster analysis process the variable weighting clustering has been significant examine subject [23-24]. In this case, it robotically calculates weight for each variable and detects the imperative variables and certain unimportant variables through variable weight. There are variables of two levels available in the data with multi view. In clustering method multi view data, the comparison of views and major role of individual variables in each view added to the dataset. In variable weighting clustering methodology, compute only the weights of individual variables and eliminate the comparison of multi view data. For that reason, these two types are not suitable for multi view data. The major difficulties of this type are these methods are not applied to the temporal datasets and also the cluster ensemble algorithm process does not consider the above mentioned two methods.

At current situation to conquer the disadvantage of the temporal data clustering representation, a method to temporal data clustering with dissimilar illustrations is used. This future work having preliminary clustering analysis on various representations to generate a multiple partitioning process and clustering ensemble algorithm is used to generate a final partition process by grouping those partitions to gather initial clustering process. At the same time in the initial clustering analysis can also be used to the previous clustering algorithms and intend a new weighted clustering ensemble algorithm with two processing phases. In the ATBCWCE structure, a weighting function consensus adjusting the input partition to person concurrence partitions based on different clustering validation. In the multi view temporal data for estimating the weight the Big Bang big Crunch is implemented the Automated TW-k means Weighting Clustering Ensemble Structure. It is known as ATBCWCE. The common analysis has also been done by using the ATBCWCE algorithm. It also exhibit effectiveness and efficiency of this method for different kind of the temporal data clustering method and also it is very easy to use for the entire internal factors are fixed in the simulation process.

Motivation and Problem Specification

In the classification method the ensemble algorithm is used [25], ensemble methodology is also used in a pattern categorization and data mining process too. Clustering ensemble algorithm is used to grouping the multiple partitioning of the data to get a single clustering outcome. There are two phase available in the clustering ensemble algorithm. The first phase is producing a number of clustering result and the second phase is known as ensemble creation, at the same time the second phase is also known as clustering ensemble trouble. In common, individual clustering results are produced from different perspectives, like various data sample subsets [26].

Fred and Jain [27] was implemented an Evidence Accumulate Clustering (EAC) algorithms. It is an instance based clustering ensemble method. Normalized similarity matrix using a EAC algorithms and its ij^{th} unit accumulates the number of decreasing data I and data j into a similar partition all the way through the primary clustering ensemble algorithms. Certainly, the primary clustering ensemble matrix maps to a novel feature space and it is a conceptual of the real ensembles. It is known as co-association matrix. After that a hierarchical based clustering algorithm such as Single Linkage (SA) can be used to extract the consensus clustering from the co-association space. The outcome of this methodology demonstrates the grouping of simple and fragile clustering methodology like k-means methodology. It represents a simple structure on any other datasets and it can reveal the innovation of the real beneath clusters with arbitrary spaces, sizes and densities. Now the four partitioning outcome are can be the result of the four various k-means algorithm.

The above mentioned partitioning methodology is not suitable for the multi view temporal data. It is the outcome of integration of multiple types of dimensions on annotations from various point of view and dissimilar types of estimation can be measured as various views. For example, the blood cell data can be partitioned into views of density, geometry, colour and texture. Each and every one is representing a view of the significant evaluation on the blood cell. In the previous method, the multi view data are used in the multi view clustering [28-30].Compared to the other clustering method this method is used to get the multiple views as a set of variables and eliminate a dissimilar content between the various views. And the multi view clustering reveals the data's from multiple views and get the dissimilar content between the various views into the appropriate to generate a exact result and partitioning of the data. This multi view data having two levels of variables. In the clustering the multi view data, comparisons of views and the role of individual variables in every view are stored in the dataset.

But the other algorithms in [28-30] have major difficulties that these algorithms are not suitable for the temporal datasets. To overcome these problems the TW-k-means methodology are proposed and also the two level variable weighting k-means clustering algorithms for the multi view data. The TW-k-means algorithms [31] is used to compare the impacts of various views and various variables in the clustering process, the weights of views and individual variables are implemented to the distance function.

The temporal clustering analysis process gives the essential path to determine the inherent structure and compact data in excess of temporal data by reveal changeable characteristics beneath temporal data in an unauthorized path. In common, in the clustering analysis process there are two troubles is available. One is model selection and another one is grouping. Some of the inherent clusters are not covered underlying temporal dataset, at the same time the coherent sequences are combined together to create a cluster matching. Specifically, present experiential topic [26] exposes that temporal data clustering focus a major trouble in temporal data mining owing to the high dimensionality and complex temporal correlation. For that only the TW-k-means algorithm is introduced to overcome this trouble. However, in the TW-k-means methodology the estimation of weight values and the clustering ensemble methodology is not suitable for all the work. For that reason, the Big Bang Big Crunch is implemented to calculate the weight value in the Automated TW-k means Weighting Clustering Ensemble structure for multi view temporal data namely ATBCWCE.

Proposed Methodology

Cluster ensemble algorithm produce numerous of various clustering and grouped together automatically and the exact consensus clustering. Specifically, present experiential topic [26] exposes that temporal data clustering focus a trouble on the temporal data mining owing to the high dimensionality and complex temporal correlation. However, in the TW-k-means methodology the estimation of weight values and the clustering ensemble methodology is not suitable for all the work. For that reason, the Big Bang Big Crunch is implemented to calculate the weight value in the Automated TW-k means Weighting Clustering Ensemble structure for multi view temporal data namely ATBCWCE. At first, a simple clustering ensemble structure to demonstrate the compassion of the ATBCWCE algorithm.

Temporal data representations

The temporal data can be written as:

$$\{x(t)\}_{t=1}^{T}$$

In this equation explains with a length of T temporal data points, using two representation namely Piecewise Local Statistics (PLS) and Piecewise Discrete Wavelet Transform (PDWT) are implemented in the previous methods [32] and also the two classical global representation are available namely Polynomial Curve Fitting (PCF) and Discrete Fourier Transforms (DFTs).

PLS representation

A porthole of the preset size is used to block time series in to set of pattern. For every pattern, the first and second order statistics are used as features of the pattern. For segment n, its local statistics μ_n and σ_n are estimated by

$$\mu_n = \frac{1}{|w|} \sum_{t=1+(n-1)|W|}^{n|W|} x(t) \tag{1}$$

$$\sigma_n = \sqrt{\frac{1}{|w|} \sum_{t=1+(n-1)|W|}^{n|W|} [x\,(t) - \mu_n]^2} \tag{2}$$

where |W| is the size of the window.

PDWT representation

At the proper levels, the Discrete Wavelet Transform (DTW) is applied to decay every segment through the low pass and high pass filtering. At level j, high-pass filters Ψ_H^jdetermine the thorough data, at the same time as low pass filters Ψ_L^jtypifycrudedata. For the nth segment, a multi scale scrutiny of J levels leads to a local demonstration with all coefficients:

$$\{x(t)\}_{t=(n-1)|W|}^{n|W|} \Rightarrow \left\{\{\Psi_L^j, \{\Psi_H^j\}_{j=1}^{J}\right\} \tag{3}$$

On the other hand, the dimension of this representation is the window size |W|. The similar mapping methods are applied to the mapping wavelet coefficients nonlinearly onto a specific low dimensional space to create a PDWT representation for decrease the dimensionality.

PCF representation

The time series is used to the parametric polynomial function [33]. It can be represented as follows:

$$x(t) = \alpha_R t^R + \alpha_{R-1} t^{R-1} + \cdots + \alpha_1 t + \alpha_0 \tag{4}$$

In the above equation the polynomial coefficients are represented as$\alpha_r (r = 0, \dots R)$ of the Rth order. All $R + 1$ coefficients obtained via the optimization constitute a PCF representation, a location-dependent global representation.

DFT representation

Subsequently, to determine the global representation of the time series the Discrete Fourier Transform method is used [34]. The DFT of time series $\{x(t)\}_{t=1}^{T}$ yields a set of Fourier coefficients:

$$a_d = \frac{1}{T}\sum_{t=1}^{T} x(t)\exp\left(\frac{-j2\pi dt}{T}\right), d = 0,1 \ldots T-1 \quad (5)$$

After that, maintain only some sample coefficients for automatically in opposition to the noise d ($d \ll T$). It means the real and imaginary parts, according to the low frequencies gathered to create a Fourier form, a position-independent global representation.

Weighted consensus function

The common thought of the weighted consensus function is the used to get the similarity between the objects in a partition for evident accumulation. This similarity matrix is determined from the weighted partitions and weights are calculated by measuring the clustering quality with various clustering process. after that, a clustering is implemented based on the entire similarity matrices to produce a person consensus partitions.

Partition weighting scheme

Assume that $X = \{x_n\}_{n=1}^{N}$ is a data set of n=1 to N objects and there are M= 1 to m partitions $P = \{P_m\}_{m=1}^{M}$, where the cluster number in M partitions could be vary. This information is gathered from the initial clustering analysis. Partition weighting method fixes a weight w_m^{π} to everyP_m in terms of a clustering validation criterion, and weights for the entire partitions based on the criterion selectively form a weight vector $w = \{w_m^{\pi}\}_{m=1}^{M}$ for the partition collection P. In the partition weighting method, describe a weight:

$$w_m^{\pi} = \frac{\pi(P_m)}{\sum_{m=1}^{M}\pi(P_m)} \quad (6)$$

where $w_m^{\pi} > 0$ and $\sum_{m=1}^{M} w_m^{\pi} = 1$. $\pi(P_m)$Is the clustering validity index value in term of the criterion π . instinctively, the weight of a partition would express its contribution to the grouping in terms of its clustering quality measured by the clustering validation criterion π.

In the above equation (6), assume the weight values depending on single view only not for multi view so it is comprehensive to multi view data clustering.

Figure 1 demonstrates the ATBCWCE structure. The set of partition matrix M with N objects represented by the set A of variables $P = (P_{1n}, P_{2n}, \ldots . P_{mn})$. Consider A is partitioned to T views $\{G_t\}_{t=1}^{T}$. In this case, $G_t \cap G_s = \emptyset$ for $s \neq t$ and $\bigcup_{t=1}^{T} G_t = A$. Allow $W = \{w_1, w_2, \ldots w_T\}$ be a set of T view weights. In this case, w_t is referred to as weight and this value is assigned to the tth view and$\sum_{t=1}^{T} w_t = 1$. In the set of m variable weights allow$V = \{V_j\}$. In this case, v_j is referred to as the weight and it is assigned to the jth variable and $\sum_{j\in G_t} v_j = 1, (1 \le t \le T), \sum_{j=1}^{o} v_j = T$. Consider k cluster available in X. determine the set of k clusters from G. determine the set of k cluster from G. and also validate the views from the view weight matrix $W = \{w_1, .. w_T\}$ and validate the variable from the variable weight matrix $V = [v_j]_o$. suppose the set of individual variables in data X considered as G and this trouble is corresponding to the individual variable weighting. Consequently, the two level variable weighting methodology as a simplification of the present variable weighting methods.

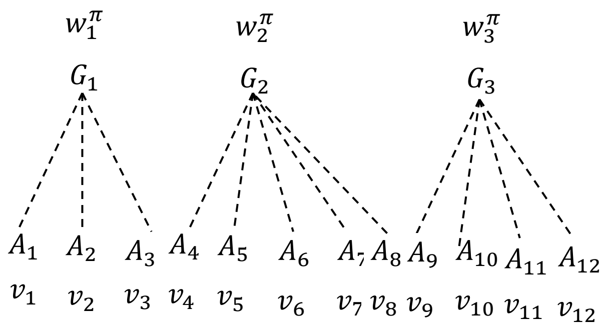


Figure 1: Example of ATBCWCE framework

In the two levels variable weighting methodologies, the variable weights V are used to validate the specific variables in every view and the view weights W are used to validate the compact cluster structure. If the views having the condensed cluster structures, a big view weight is applied so as to improve the effect of the view. On the divergent, if the view having simple cluster structures, a minute view weight is applied to reduce the effect of the view.

Problem formulation

The clustering process to partition P_{mn} into k clusters with weights for both views and individual variables is modelled as minimization of the following objective function:

$$F(U,Z,V,W) = \sum_{i=1}^{m}\sum_{j=1}^{o}\sum_{t=1}^{T}\sum_{l=1}^{k} u_{i,l}\, w_t v_j d(P_{i,j}, z_{l,j}) + \eta\sum_{j=1}^{o} v_j \log(v_j) + \lambda\sum_{t=1}^{T} w_t \log(w_t) \quad (7)$$

Focus to

$$\begin{cases} \sum_{l=1}^{p} u_{i,l} = 1, u_{i,l} \in \{0,1\}, 1 \le i \le m \\ \sum_{t=1}^{T} w_t = 1, 0 \le w_t \le 1 \\ \sum_{j=1}^{o} v_j = 1, 0 \le w_t \le 1, 1 \le t \le T \end{cases} \quad (8)$$

Where
U is referred to as $m \times k$ partition matrix whose essentials$u_{i,l}$ are binary where $u_{i,l} = 1$ indicates that object i is owed to cluster l

$Z = \{Z_1, Z_2, \dots Z_k\}$s a set of k vectors demonstrating the centres of the k clusters;

$W = \{w_1, w_2, \dots w_T\}$are T weights for T views;
$V = \{v_1, v_2, \dots, v_o\}$are o weights for o variables;
$\lambda > 0, \eta > 0$ are two given parameters;
$d(P_{i,j}, z_{l,j})$ is a distance or dissimilarity measure on the j^{th} variable among the i^{th} object and the centre of the l^{th} cluster. If the variable is arithmetical, then:

$$d\left(P_{i,j}, z_{l,j}\right) = \left(P_{i,j} - z_{l,j}\right)^2 \tag{9}$$

If the variable is categorical, then:

$$d\left(P_{i,j}, z_{l,j}\right) = \begin{cases} 0, (P_{i,j} = z_{l,j}) \\ 1, (P_{i,j} \neq z_{l,j}) \end{cases} \tag{10}$$

In the first equation is estimating the addition of the cluster dispersions. The second and third equation is the two negative weight entropies. For the more number of views and variables to manage the incentive for clustering process two positive parameters λ and η is used. There are four steps are used to solve the minimization trouble by iteratively based on (7).

1. Problem Prb$_1$: Fix $Z = \widehat{Z}, V = \widehat{V}$ and $W = \widehat{W}$, and solve the reduced problem $P(U, \widehat{Z}, \widehat{V}, \widehat{W})$;
2. Problem Prb$_2$: Fix $U = \widehat{U}, V = \widehat{V}$, and $= \widehat{W}$, and solve the reduced problem $P\left(\widehat{U}, Z, \widehat{V}, \widehat{W}\right)$;
3. Problem Prb$_3$: Fix $U = \widehat{U}, Z = \widehat{Z}$ and $= \widehat{W}$, and solve the reduced problem $P\left(\widehat{U}, \widehat{Z}, V, \widehat{W}\right)$;
4. Problem Prb$_4$: Fix $U = \widehat{U}, Z = \widehat{Z}$, and $V = \widehat{V}$, and solve the reduced problem $P\left(\widehat{U}, \widehat{Z}, \widehat{V}, W\right)$;

Let $U = \widehat{U}, Z = \widehat{Z}$, and $W = \widehat{W}$ be fixed. $P\left(\widehat{U}, \widehat{Z}, V, \widehat{W}\right)$is minimized if and only if

$$v_j = \frac{\exp\left\{\frac{-E_j}{\eta}\right\}}{\sum_{h \in G_t} \exp\left\{\frac{-E_h}{\eta}\right\}} \tag{11}$$

$$E_j = \sum_{l=1}^{k} \sum_{i=1}^{m} \hat{u}_{i,l} \widehat{w}_t \, d\left(P_{i,j}, \hat{z}_{l,j}\right) \tag{12}$$

In the above equations t is referred to as the index of the view that the j^{th} variable is assigned to the process.

Let $U = \widehat{U}, Z = \widehat{Z}$, and $V = \widehat{V}$ be fixed. $P\left(\widehat{U}, \widehat{Z}, \widehat{V}, W\right)$is minimized if and only if

$$w_t = \frac{\exp\left\{\frac{-D_t}{\eta}\right\}}{\sum_{h=1}^{T} \exp\left\{\frac{-D_h}{\eta}\right\}} \tag{13}$$

$$D_t = \sum_{l=1}^{k} \sum_{i=1}^{m} \sum_{j \in G_t} \hat{u}_{i,l} \hat{v}_j \, d\left(P_{i,j}, \hat{z}_{l,j}\right) \tag{14}$$

To manage the allocation of the two types of weights V and W the input parameters λ and η are used. If $\lambda = 0$ and $\eta = 0$ means it is very easy to validate the objective functions with respect to the V and W. Furthermore, they are used as follows:

- η> 0. In this case, according to (11), v is contrariwise comparative to E. The smaller E_j, the larger v_j, the more significant the equivalent variable.
- η = 0. In this case, according to (11), $\eta = 0$ will generate a clustering outcome with only one important variable in a view. It may not be enviable for high dimensional data.
- λ> 0. In this case, according to (14), w is contrariwise comparative to D. The smaller D^t, the larger w_t, the more condensed the equivalent view.
- λ = 0. In this case, according to (14), $\lambda = 0$ will generate a clustering outcome with only one important view. It may not be enviable for multi view data.

In common, λ and η are set as positive real values.

Big Bang–Big Crunch (BB–BC)

Randomness can be seen as corresponding to the energy debauchery at the same time as convergence to a local or global optimum point. This is known as gravitational attraction. The GA and BB-BC methodologies are similar and both are used to generate an initial population arbitrarily. In this proposed work the weights of views and individual variables are considered as inputs to the k cluster. The generation of initial population arbitrarily from the weights values of the view and individual variables is known as Big Bang phase. In these particular steps, the weight values of the persons are broadening the entire search space unvaryingly. The Big Bang phase method is working on the basis of Big Crunch phase. The Big Crunch is one of the convergence operators that have more number of weights for views and individual variables as input but only one clustering outcome can be generated beneath k clusters and it is known as centre of mass. By using this centre of mass value the cluster can be estimated. At this point, the centre of mass refers to the opposite of the fitness function value.

The following formula is used to estimate the optimal weight calculation and it is represented by x_c

$$x_c = \frac{\sum_{i=1}^{n} \frac{1}{f^i} x_i}{\sum_{i=1}^{n} \frac{1}{f^i}} \tag{15}$$

In the above equation x_i is referred to as the selected weights for both views and individual variables surrounded by an n-dimensional search space produced, the fitness value of the data is represented by f^i, N is represented as the population size in Big Bang phase. To evaluate the centre of gravity the convergence operator is used by choosing the two dissimilar weight values. In the Big-Crunch phase both the views and individual variables weights are selected as members and it is used as a contraction operator. After this Big Crunch phase process, the novel weights for both views and individual variables are generated. Subsequently this value can be used the next iteration process.

The second detonation subsequent to, the fitness value of every centre and cluster data point is re-estimated. These consecutive detonation and reduction steps are conceded continuously in anticipation of a stopping criterion has been met. For both the views and individual variables the elapsed iteration generates good weights and reduces a value for the partition. In the iteration process, the partition will reach zero value means this partition process goes to infinity stage. The following steps are followed in BB–BC method:

Step 1 for both individual variables and views the initial generation as number of weights can be formed with N persons in a random manner.

Step 2 by using equation (11) & (14) to evaluate the fitness function values of the whole weights for both views and individual variables.

Step 3 by using the equation (15) the centre of mass was calculated. After that the final weight values are chosen.

Step 4 Estimate the weight values in the region of the fitness function by totalling or subtracting a normal random amount whose value losses the iterations intervene. This can be written as follows

$$x^{new} = x^c + \frac{lr}{k} \qquad (16)$$

In the above equation $x^c$is represented as centre of mass, l is represented as the upper limit of the factors, normal random is represented as r and the iteration step can be referred to as k. Then novel point x^{new} is an upper and lower bounded.

Step 5 Return to Step 2 untilcertain stopping criterion has been met.

In terms of various kinds of partitions a weighted similarity matrix S is used to replicate the collective relationship between the entire data. This matrix is usually tended to the collect confirmation for the clustering quality and for this reason the entire partitions are treated in different manner. For sturdiness next to noise, do not grouping three weight similarity matrices openly, but practice to yields three candidates' consensus partitions.

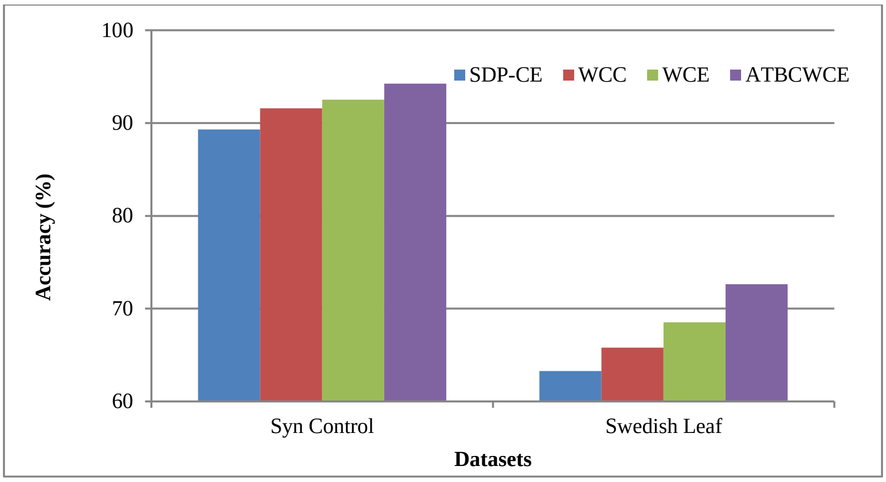
Experimentation Results

For assessment purpose, the ATBCWCE methodology is applied for gathering the time-series benchmarks for temporal data mining [35]. To estimate a time series categorization and clustering algorithms in context to temporal data mining there are sixteen synthetic or time series datasets are gathered [35]. In this anthology, by using the already available information like the number of classes K and the time series class label in a data set auxiliary splits into training and testing subsets. This training and testing subsets are used to estimate the classification algorithm. In this process, the entire sixteen datasets are used. Both the training and test subsets are available in this datasets. The standard temporal data clustering procedures are openly applied on the time series to extract their recital on benchmark collection. This is attained by model based algorithm and temporal proximity.

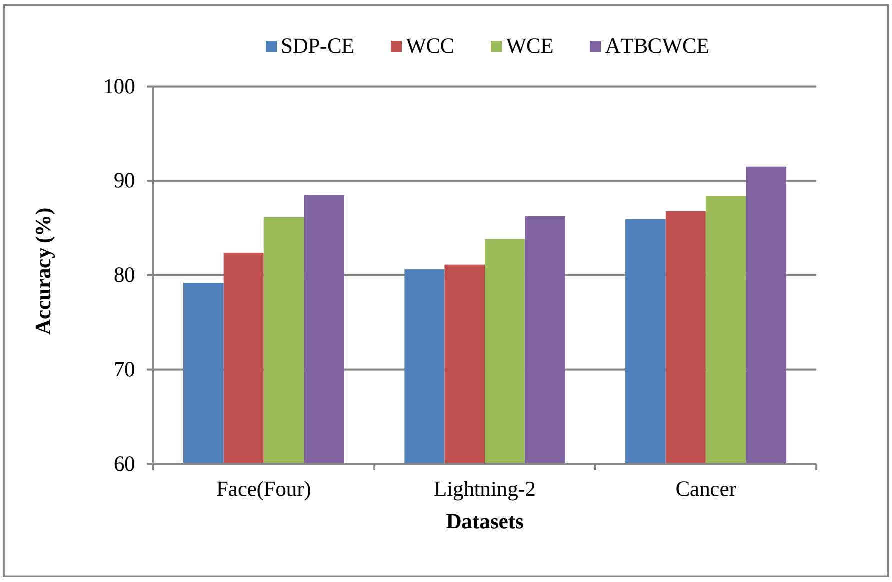
For the comparing purpose, make use of the three state algorithms implemented from various points of view. It means Semi Definite Programming based Clustering Ensemble (SDP-CE) algorithm [20], Weighted Clustering Ensemble (WCE) algorithm [37], Weighted Consensus Clustering (WCC) algorithm [36] and ATBCWCE algorithm. In this work, for the first clustering analyzing process the k-mean step is used. By using the given information the three algorithms are implemented without addressing the model selection. And in the k-mean step the accurate cluster number of every data set $K^*$is used. In the process there are ten partitions are produced for a single demonstration. Therefore, totally 40 partitions are produced and these partitions are grouped together for every dataset. These similar methods are used in the WCE also for k-mean to generate a partition. And the K values are chooses arbitrarily from$K^* - 2 \leq K \leq K + 2(K > 0)$. After the twenty trails, the standard deviation and average of classification precision rates is stored in the dataset. In this work, the four dataset samples are used for the clustering process namely Syn Control, Face (four), Swedish Leaf, Two patterns and Lightning-2.

Table 2 illustrate the collective list of each and every outcome of the clustering methods. First of all for the clustering process the accurate cluster number for four datasets have been calculated. This cluster number was mentioned in Table 2. It indicates the clustering temporal data of high dimension. Specifically, the SDP-CE algorithm took a more time to complete a comparison process with other algorithms including the clustering ensemble algorithm. The ATBCWCE algorithms produce a better outcome when contrast to certain algorithms. From observing the Table 2, the ATBCWCE is usually better than the SDP-CE algorithm on the suitable representation space and the preeminent parameter setup.

For example, considering the Swedish Leaf dataset the ATBCWCE produce 72.65% outcome correctness. When this outcome is compared to the SDP-CE, WCC and WCE methods it is 9.37%, 6.84% and 4.13% higher respectively. The final outcome of five different dataset and their clustering methods are depicted in Figure 2 (a) and (b).



(a) Accuracy of different clustering methods to Syn Control and Swedish Leaf



(b) Accuracy of different clustering methods to Face(Four) , Lightning-2 and Cancer
Figure 2: Clustering Precision (in Percent) of Diverse Clustering Algorithms on Time-Series Benchmarks

Table 1
Time-Series Benchmark Information [35]

Dataset	Number of class (K*)	Size of dataset (Training +testing)	Length
Syn Control	6	300+300	60
Gun-Point	2	50+150	150
CBF	3	30+900	128
Face(all)	14	560+1690	131
OSU leaf	6	200+242	427
Swedish Leaf	15	500+625	128
50 words	50	450+455	270
Trace	4	100+100	275
Two patterns	4	1000+4000	128
Wafer	2	1000+6174	152
Face(Four)	4	24+88	350
Lightning-2	2	60+61	637
Lightning-7	7	70+73	319
ECG	2	100+100	96
Adiac	37	390+391	176
Yoga	2	300+3000	426
Cancer	10	2100+5100	7100

Table 2
Clustering Accuracy (in Percent) of Different Clustering Algorithms on Time-Series Benchmarks [35]

Dataset	Accuracy (%)				Error (%)			
	Different representations				Different representations			
	SDP-CE	WCC	WCE	ATBCWCE	SDP-CE	WCC	WCE	ATBCWCE
Syn Control	89.32	91.58	92.51	94.26	10.68	8.42	7.49	5.74
Swedish Leaf	63.28	65.81	68.52	72.65	36.72	34.19	31.48	27.35
Face(Four)	79.21	82.38	86.14	88.53	20.79	17.62	13.86	11.47
Lightning-2	80.63	81.14	83.85	86.26	19.37	18.86	16.15	13.74
Cancer	85.96	86.78	88.41	91.51	14.04	13.22	11.59	8.49

For example, considering the cancer dataset samples proposed ATBCWCE produces 91.51% accuracy which is higher when compared to other existing algorithms.

By measuring the clustering precision cast-off Precision, F-measure, Recall, and average cluster entropy, accuracy to estimate the outcomes. This can be described as follows:
Precision: It is used to estimate the fraction of the accurate objects in the middle of those that the algorithms to the appropriate cluster.

Recall: It is used to identify the fraction of the authentic objects.

F-measure: F-measure represents the vocal accuracy mean, recall and correctness is the quantity of precisely clustered objects.

Average Cluster Entropy (ACE) based contamination of a cluster data set. If$p_{ij}$is represented as the fraction of class j in obtained clusteri, N_i is represented as the cluster sizei, and the total amount of instances are represented byN. The average cluster entropy is definite as follows:

$$E = \sum_{i=1}^{K} \frac{N_i(-\sum_j p_{ij} \log(p_{ij}))}{N} \quad (17)$$

Where, K represents the number of clusters.

By using the above mentioned four parameters the final outcome of the clustering method are given in Table 3. In contrast to the each algorithm the ATBCWCE algorithm attains higher result in the parameter because the centroid values and weight are estimated automatically as an alternative with the fixed value.

The total clustering outcome is shown in Table 3. According this outcome, the ATBCWCE considerably to another three algorithms in nearly too every outcome, specifically on the sample from the five data set. Even though ATBCWCE is the elaboration to TW –k means and WCE methodology, weights with views on increased outcome of this process. In every five data sets the SDP-CE generate a reasonable outcome of this process. One of the very imperative scrutinise is the estimation of the clustering fault by beneath the result into different clusters. At last, in the ATBCWCEW process generate a smaller amount of clustering fault when compare to the previous methodologies because this ATBCWCE process is without any difficulties it is relevant to the temporal and multi view data.

When compared to the SDP-CE, WCC and WCE methodologies the ATBCWCE with multi view and temporal data achieve a better result. The result is 92.18% clustering precision which is 25.6%, 3.55% and 1.43% increased with respect to the data set samples.

When compared to the SDP-CE, WCC and WCE methodologies the ATBCWCE with multi view and temporal data achieve a better result. The result is 92.84% clustering precision evoke value which is 3.32%, 2.59% and 1.28% increased with respect to the data set samples.

When compared to the SDP-CE, WCC and WCE methodologies the ATBCWCE with multi view and temporal data achieve a better result. The result is 92. 5% F-measure which is 4.48%, 3.07% and 1.35% increased relating to the data set samples.

Conclusion and Future work

Here, a temporal data clustering method through a Weighted Clustering Ensemble (WCE) on different representations and additionally suggest a suitable estimation to comprehend Clustering ensemble procedures. This process is mainly spotlights on analysis of official Clustering algorithms. Moreover, the Big bang big crunch is used to calculating the WCE weight and it is the conservatory of Automated TW-k means Weighting Clustering Ensemble structure and this value is functional to the multi view temporal data. This process is known as ATBCWCE structure. This work describes ATBCWCE structure to cluster the cancer genomics data repositories, along with cluster subtypes to cluster and analyse these data. This ATBCWCE method is used to estimate the weights for the both views and individual variable from the multiple view temporal data at the same time in the clustering process. By using the compact views, both the weight values and also significant variables can be detected and it is also eliminate the noise variables, low quality views. In this proposed work the ATBC WCE calculate the multiple view weights and these weights is frankly produce a novel weighting method beneath a data set in common. This process is also exposes the convergence possessions of the view weights. When compared to the three clustering algorithms and ATBCWCE on five temporal data sets and the result revealed that

ATBCWCE significantly for another three clustering algorithms in the four assessment indices like Recall, F-measure, Accuracy and Precision. In upcoming work, the two-level variable weighting method and another method like fuzzy methodology, semi-supervised methodology and subspace clustering methodology are grouped together automatically.

Table 3
Summary of Diverse Clustering Algorithms on five dataset by four parameters

Dataset	Evaluation	SDP-CE	WCC	WCE	ATBCWCE
Syn Control	**Precision (%)**	86.58	88.63	90.75	92.18
	Recall (%)	89.52	90.25	91.56	92.84
	F measure (%)	88.02	89.43	91.15	92.50
	ACE	2.85	1.92	1.58	1.23
Swedish Leaf	**Precision (%)**	67.53	68.81	69.71	72.95
	Recall (%)	69.82	71.25	72.83	74.18
	F measure (%)	68.65	70.01	71.24	73.56
	ACE	1.57	1.28	1.15	0.76
Face(Four)	**Precision (%)**	80.23	81.56	84.19	86.52
	Recall (%)	81.29	82.53	84.23	87.46
	F measure (%)	80.76	82.04	84.21	86.98
	ACE	1.63	1.42	1.38	1.26
Lightning-2	**Precision (%)**	80.52	81.47	82.58	83.79
	Recall (%)	81.26	82.18	83.45	84.15
	F measure (%)	80.88	81.82	83.01	83.96
	ACE	1.93	1.67	1.28	1.09
Cancer	**Precision (%)**	86.93	87.53	88.19	91.58
	Recall (%)	89.32	91.51	92.37	93.81
	F measure (%)	88.125	89.52	90.28	92.695
	ACE	1.23	1.15	1.08	0.95

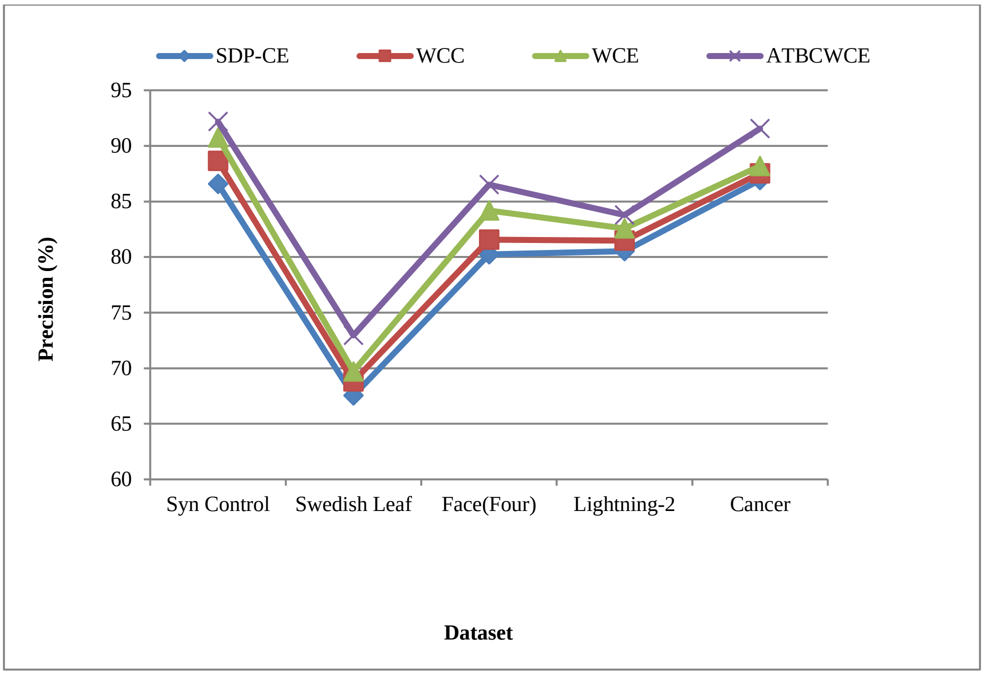


Figure 3: Precision of Different Clustering Algorithms on Time-Series Benchmarks datasets

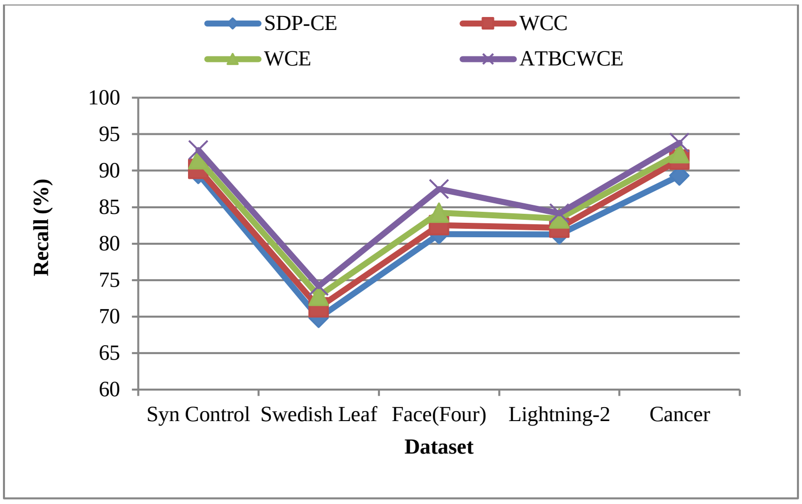


Figure 4: Recall of Different Clustering Algorithms on Time-Series Benchmarks Datasets

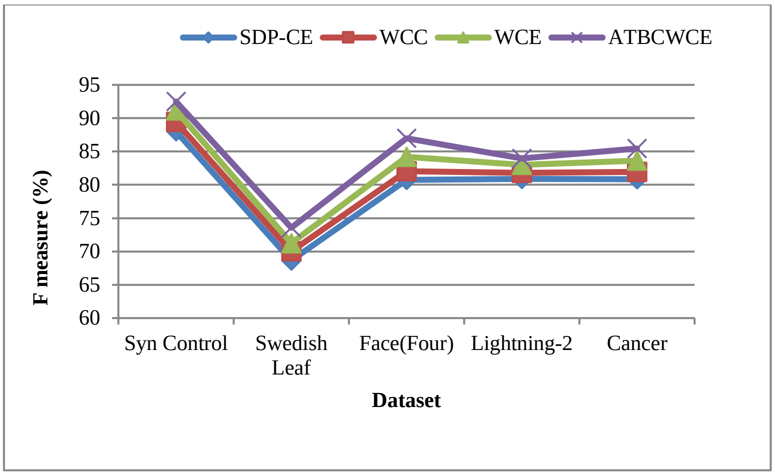


Figure 5: F-measure of Different Clustering Algorithms on Time-Series Benchmarks datasets

References

1. Arthur, D., & Vassilvitskii S. (2007). K-means++: the advantages of careful seeding, in: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, pp. 1027–1035.

2. Cuturi, M., (2011). Fast global alignment kernels, in: Proceedings of the 28th International Conference on Machine Learning (ICML-11), pp. 929–936.

3. Cuturi M., Vert, J.P., Birkenes, Ø., & Matsui T. (2007). A kernel for time series based on global alignments, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing. 11, pp. 413–416.

4. Petitjean, F., Ketterlin, A., & Gançarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering, Pattern Recognition. 44 (3), 678–693.

5. Dhillon, I. S., Guan, Y., & Kulis, B. (2004). Kernel k-means: spectral clustering and normalized cuts, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, pp. 551–556.

6. Liao, W. (2005). Clustering of time series data – a survey. Pattern Recognit. 38, 1857–1874.

7. Frambourg, C., Douzal-Chouakria, A., & Gaussier, E. (2013). Learning multiple temporal matching for time series classification, in: A. Tucker, F. Höppner, A. Siebes, S. Swift (Eds.), Intelligent Data Analysis, Springer Berlin Heidelberg, London, pp. 198–209.

8. Liao, T. W. (2007). A clustering procedure for exploratory mining of vector time series. Pattern Recognition, 4(9), 2550-2562.

9. Dacheng, N., Yan, F., Junlin, Z., Yuke, F., & Hu, X. (2010).Time series analysis based on enhanced NLCS, 3rd International Conference on Information Sciences and Interaction Sciences (ICIS), pp.292-295.

10. Wu, X., & Huang, D. (2010). Data stream clustering for stock data analysis, 2nd International Conference on Industrial and Information Systems (IIS). 2, 168-171.

11. Huang, X., & LI, H. l,. (2010). Research on Predicting Agricultural Drought Based on Fuzzy Set and RIS Analysis Model, 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), pp. 186-189.

12. Lin, Y., & Yang, Y. (2009). Stock markets forecasting based on fuzzy time series model, IEEE International Conference on Intelligent Computing and Intelligent Systems. ICIS, 1, 782-786.

13. Gao, S., He, Y., & Chen, H. (2009). Wind speed forecast for wind farms based on ARMA-ARCH model. International Conference on Sustainable Power Generation and Supply, SUPERGEN'09, 1-4.

14. Jixue, D. (2009). Data Mining of Time Series Based on Wave Cluster. International Forum on Information Technology and Applications, IFITA, 1, 697-699.

15. Powell, N., Foo, S. Y., Weatherspoon, M. (2008). Supervised and Unsupervised Methods for Stock Trend Forecasting, 40th south eastern Symposium on System Theory, SSST. pp. 203-205.

16. Wu, J. X., & Wei, J. L. (2007). Combining ICA with SVR for prediction of finance time series, Proceedings of the IEEE International Conference on Automation and Logistics, Jinan, China, pp. 95-100.

17. Verdoolaege, G., & Rosseel, Y.(2010). Activation Detection In Event-Related FMRI Through Clustering Of wavelet Distributions, Proceedings of 2010 IEEE 17th International Conference on Image Processing, Hong Kong, pp. 4393-4395.

18. Fern, X., & Brodley, C. (2004). Solving Cluster Ensemble Problem by Bipartite Graph Partitioning, Proc. International Conf. Machine Learning, pp. 36-43.

19. Gionis, A., Mannila, H., & Tsaparas, P. (2007). Clustering Aggregation, ACM Trans. Knowledge Discovery from Data, 1(1), 1-30.

20. Singh, V., Mukerjee, L., Peng, J., & Xu, J. (2007). Ensemble Clustering Using Semi definite Programming, Advances in Neural Information Processing Systems, pp. 1353-1360.

21. Topchy, A., Law, M., Jain, A., & Fred, A. (2004) Analysis of Consensus Partition in Cluster Ensemble, Proc. IEEE Int'l Conf. Data Mining, pp. 225-232.

22. Chaudhuri, K., Kakade, S., Livescu, K., & Sridharan, K. (2009). Multiview Clustering via Canonical Correlation Analysis, Proc. 26th Ann. Int'l Conf. Machine Learning, pp. 129-136.

23. Tsai, C. Y. & Chiu, C. C. (2008). Developing a Feature Weight Self-Adjustment Mechanism for a k-Means Clustering Algorithm, Computational Statistics and Data Analysis, 52(10), 4658- 4672.

24. Deng, Z., Choi, K., Chung, F., & Wang, S. (2010). Enhanced Soft Subspace Clustering Integrating Within-Cluster and Between-Cluster Information, Pattern Recognition, 43(3), 767- 781.

25. Parvin, H., Minaei-Bidgoli, B., Parvin, S., & Alinejad, H. (2013). A New Classifier ensemble methodology based on subspace learning. Journal of Experimental & Theoretical Artificial Intelligence. 25(2), 227-250.

26. Parvin, H., Minaei-Bidgoli, B.,& Parvin, S., Alinejad, H. (2013). Data weighing mechanisms for clustering ensembles. Comput Electr Eng, 39(5), 1433–1450.

27. Fred, A., & Jain, A. K. (2002). Data clustering using evidence accumulation. In: Proceedings of the 16th international conference on pattern recognition, pp.276–280.

28. Tzortzis, G., & Likas, C. (2010). Multiple View Clustering Using a Weighted Combination of Exemplar-Based Mixture Models, IEEE Trans. Neural Networks, 21(12), 1925-1938.

29. Long, B., Yu, P., & Zhang Z., (2008). A General Model for Multiple View Unsupervised Learning, Proc. Eighth SIAM Int'l Conf. Data Mining (SDM '08), pp.1-12.

30. Greene, D., & Cunningham, P. (2009). A Matrix Factorization Approach for Integrating Multiple Data Views, Proc. European Conf. Machine Learning and Knowledge Discovery in Databases, pp. 423-438.

31. Chen, X., Xu, X., Huang, J.Z., & Ye, Y. (2013). TW-k-means: automated two-level variable weighting clustering algorithm for multi view data. IEEE Transactions on Knowledge and Data Engineering, 25(4), 932-944.

32. Wang, S., & Chen, K. (2007). Ensemble Learning with Active Data Selection for Semi-Supervised Pattern Classification, Proc. Int'l Joint Conf. Neural Networks, pp. 355-360.

33. Chen, W., & Chang, S. (2000). Motion Trajectory Matching of Video Objects, Proc. SPIE/IS&T Conf. Storage and Retrieval for Media Database, pp. 544-553.

34. Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast Subsequence Matching in Time-Series Databases," Proc. ACM SIGMOD, pp. 419-429.

35. Keogh, E., (2010) Temporal Data Mining Benchmarks.

36. H. Kien, A. Hua, and K. Vu, (2008). Constrained Locally Weighted Clustering. Proc. ACM Int'l Conf. Very Large Data Bases (VLDB), pp. 90-101.

37. Rajan, C, Geetha, K, (2016), "Automatic Colorectal Polyp Detection in Colonoscopy Video Frames", Asian Pacific Journal of Cancer Prevention, vol. 17, no 11, pp. 4869-4873.

38. Yang, Y., & Chen, K. (2011). Temporal data clustering via weighted clustering ensemble with different representations. IEEE Transactions on Knowledge and Data Engineering, 23(2), 307-320