Data Mining For Selected Clustering Algorithms: A Comparative Study

Omar Y. Alshamesti
Department of Computer science
Palestine Technical Colleges/Al-aroub
Hebron, Palestine
omar@ptca.edu.ps

Ismail M. Romi
College of Administrative sciences and Informatics
Palestine Polytechnic University
Hebron, Palestine
ismailr@ppu.edu

Abstract— ***Data mining is the process used to analyze a large amount of heterogeneous data to extract useful information from it. Clustering is one of the main data mining techniques used to divide the data into several groups and each group is called a cluster, which contains objects that are homogeneous in one cluster and different from other clusters. As a reason of many application that depend on clustering techniques, and since there is no combined method for clustering, this paper focus on the comparison between k-mean, Fuzzy c-mean, self organizing map (SOM) and support vector clustering (SVC) to show how those algorithms solve the clustering problem, compare the new methods of clustering (SVC) with the traditional clustering methods (K-mean, fuzzy c-mean and SOM), and show how the studies improve SVC algorithm. The results shows that SVC is better than the k-mean, fuzzy c-mean and SOM; because it has no dependency on either the number or shape of the cluster, it can deal with outlier, overlapping, and it depends on the kernel method. Finally this paper shows that the enhancement using the gradient decent and the proximity graph improve the support vector clustering time by decreasing its computational complexity to O(nlogn) instead of $O(n^2d)$, but the practical total time for improvement support vector clustering (iSVC) labeling method is better than the other methods that improve SVC.***

Keywords- Data Mining, Clustering, Self-Organizing Map, Support Vector Clustering, Computational Complexity.

I. INTRODUCTION

In the early 1990's, the establishment of the internet made a huge amount of data to be stored electronically; therefore, handling this amount of data became to be a necessity. For this reason, data mining emerged; it can be defined as the process used to analyze a large amount of heterogeneous data to extract useful information from it [12] using several techniques such as clustering. Clustering process [9] is an unsupervised learning technique that is used to divide the data into several groups and each group is called a cluster which contains objects that are homogeneous in that cluster and should be different from other clusters. Many clustering algorithms have been proposed by researchers [13, 14] that can be used to divide the data into several groups based on similarity. Some of these algorithms are K-mean, fuzzy c-mean, Self Organizing Map (SOM) and Support Vector Clustering (SVC). K-mean is a well known partitioning method and one of the most popular clustering algorithms used in scientific and industrial application [2]. Fuzzy c-mean [1, 17] is an iterative algorithm that is frequently used in pattern recognition and is allowed one piece of data to belong to more than one cluster. SOM algorithm proposed by Kohenen in 1982 can be classified as a powerful method for clustering high dimensional data [4]. SVC [24] is a nonparametric clustering process which depends on Support vector machine (SVM) concepts.

Many applications depend on clustering techniques, while there is no fixed method or technique to perform this process. This encourages researchers to keep developing clustering and its techniques, where many studies improve data clustering algorithms [16, 17, 26, 27, 29, 30], other studies implement new methods [18, 19, 24], and furthermore studies was compared different data clustering algorithm for different factors [1, 2, 5, 6, 7, 8, 20]. This paper will focus on comparing k-mean, fuzzy C-mean, SOM and SVC algorithms in order to show how those algorithms solve clustering problems, and then compare those traditional methods with the new clustering method; mainly SVC, in order to find out the improvements and characteristics that reduce the computational complexity of this algorithm. Those comparisons will provide a tool for selecting the best clustering algorithm in specified area such as text mining, geographical information system, and information retrieval that depend on clustering.

II. BACKGROUND AND LITERATURE REVIEW

Data mining is the process of analyzing data from different perspective and summarizing it as useful information [12]. There are many data mining techniques that can be used to analyze data such as classification and clustering [32, 13]. Those techniques are based on two type of learning paradigms [32] which are supervised learning and unsupervised learning. Clustering is one of these techniques that depened on unsupervised learning paradigm that is used to divide the data into several groups and each group is called a cluster. Many algorithms are proposed for data clustering; these algorithms can be divided into two main groups: hierarchal and

partitioned algorithms [14]. The main drawback of partitioned algorithm is the chosen of number of clusters. In an attempt to solve this problem a prior research suggest an algorithm which is called SVC. The next section will be investigated in the selected clustering algorithms and it introduces a brief idea about these algorithms.

A. Clustering technique: selected algorithms

Clustering is the process used to divide the data into several groups and each group is called a cluster. There are many clustering algorithms are proposed to solve the clustering problem. In this paper four clustering algorithms are chosen to study and compare. The algorithms that are chosen: k-mean, fuzzy c-mean, SOM and SVC.

K-mean clutsering algorithm

K-mean was invented by [Hartigan 1975, Hartigan and Wong 1979] and its name comes from representing each cluster by the mean and this is called centroid. A lot of studies in the prior research found that K-mean is a well known partitioning method and the most popular clustering algorithm used in scientific and industrial application [2]. K-mean objective function is to minimize the average squared distance of the object from their cluster center, where the cluster center is defined as the mean of the objective in a cluster C as in equation 1. The main advantages of this algorithm are fast and easy to implement [1, 2] whereas the disadvantages of this algorithm has no way to deal with outliers which is the data points that don't belong to any cluster [2].

$$\mu(c) = \frac{\sum xi}{|c|} \tag{1}$$

C is the number of clusters.

K-mean algorithm

1. Chooses the number of clusters, K.
2. Selects k points as an initial centroid of clusters.
3. Classifies each vector into the closest center by Euclidean distance measure.

$$\|xi - ci\| = \min \|\|xi - ci\|\| \tag{2}$$

4. Recomputed the cluster center as in equation 3.

$$C(i) = \frac{\sum xi}{ni} \tag{3}$$

5. If none of the cluster center changes in step 4, stop; otherwise go to step 3.

Fuzzy c-mean algorithm

Fuzzy c-mean [1,17] is an iterative algorithm that is frequently used in pattern recognition and allows one piece of data to belong to more than one cluster by a degree of membership, which define the percentage by which the data point belong to the cluster. FCM runs by finding the cluster center that minimizes the dissimilarity function as in equation 4.

$$jm = \sum_{i=1}^{n} \sum_{j=1}^{n} \bigcup_{ij}^{m} \|xi - cj\| \tag{4}$$

Where m is a real number greater than 1, *Uij* is the degree of membership of *Xi* in the cluster *J*, *Xi* is the *ith* of d-dimensional center of the cluster and $\|*\|$ is used to express the similarity between any measured data and the cluster.

Fuzzy c-mean algorithm

1) Initialize $U = [Uij]$ matrix, $U(0)$. $U(0)$
2) Calculate the center vector for each step by computing:

$$Vij = \frac{\sum_{i=1}^{n} \bigcup_{ik}^{m} * Xkj}{\sum_{i=1}^{n} \bigcup_{ik}^{m}} \tag{5}$$

3) Calculate the distance matrix by computing:

$$Dij = \sqrt{\left[\sum_{j=1}^{m} Xkj - Vij\right]} \tag{6}$$

4) Update the membership matrix (U(k), U(k+1) by computing:

$$Uij = \frac{1}{\sum_{j-1}^{c} \left[\frac{xi - cj}{xi - ck}\right]^{\frac{2}{m-1}}} \tag{7}$$

5) If $\|U(k+1) - U(k) < \varepsilon\|$ then stop otherwise return to step 2.

Self organizing map (SOM) algorithm

Self organizing map (SOM) was proposed by Chokemen in 1982. It is a powerful method for clustering high dimensional data [4]. SOM algorithm is an artificial neural network used to map the high dimensional data into low dimensional space which is usually two dimensional space called map as in Figure 1. This map consists of a number of neurons or units and each one is represented by a weight vector [4]. The Kohonen neural network consists of two layers: the input and output layer. The input layer contains the dataset vectors while the output layer forms a two dimensional array of nodes. The aim of the SOM algorithm is to put the sample unit in the map and then close together the similar sample units. The virtual units are modified iteratively through the artificial neural network (ANN) during the training process.

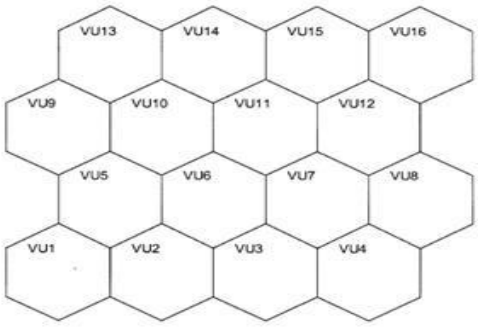

Figure1. A Self-Organizing Map formed by a rectangular Grid with a virtual unit V_{Uk} in each hexagon, [4]

SOM Algorithm

1. Initialize the virtual units using random sample drawn from the input dataset.
2. Choose a random sample unit as an input unit.
3. Compute the Euclidean distance between the sample unit and each virtual unit W_i.
4. Choose the closest virtual unit to the sample unit as a wining unit or neuron and it is called the best matching unit BMU.
5. Update the virtual unit using the following rule:

$$\omega_{ik}(t+1) = \omega_{ik}(t) + h_{ck}(t)\left[x_{ij}(t) - \omega_{ik}(t) \right] \quad (8)$$

Where t the time and h_{ck} is a neighborhood function which can be computed in several ways. The most common studies use the Gaussian function:

$$h_{ck}(t) = \exp\left(\frac{\left\| r_k - c_k \right\|^2}{2\sigma^{2(t)}}\right) \quad (9)$$

Where r_c and r_c are the position of neuron t and c in the SOM grid. And σ is the learning factor and it is a decreasing function of the time. σ converge to 0.

6. Increase the time $t = t + 1$. If $t < t_{\max}$ then go to step 2 else stop training.

Support vector clustering

SVC clustering is a nonparametric clustering algorithm that is based on support vector machine (SVM) proposed by [25]. The author in [24] proposes this clustering method that searches for the clustering solution without any assumption of their numbers or shapes.

SVC Algorithm

1. Maps the data points from a data space to high dimensional space called feature space
2. Finds the smallest sphere that encloses the image of the data.
3. Maps the sphere back to the data space.
4. The mapped sphere forms a set of contours which enclose the data point.

These set of contours that enclose the data points are interpreted as cluster boundaries. The number of cluster can be increased or decreased depending on the kernel width [24]. SVC algorithm can deal with outlier using soft margin constrains and with overlapping cluster using a large value of kernel width.

Optimization stage

As we mention above SVC method [24] transform the data space to some high dimensional spaces and then the following tasks are performed:

1. Looks for the smallest sphere that encloses a set of data point. This process is described by equation 10:

$$\left\| \Phi(xj - a \right\|^2 \le R^2 \quad (10)$$

soft constrains are employed to allow some data point to be enclosed in the sphere by adding a slack variable. This process is presented as:

$$\left\| \Phi(xj - a \right\|^2 \le R^2 + \varepsilon_j \quad (11)$$

2. Uses lagrangian multiplier introduced by [24] to solve this problem because it is an optimization problem.

$$L = (R^2 + \varepsilon_{j-} \left\| \Phi(x_j) - a^2 \right\| \beta_j \mu_j + C \sum \varepsilon_j \quad (12)$$

3. Derives the equation (12) with respect to R, then the result are:

$$\sum \beta_j = 1 \quad (13)$$

$$a = \sum \beta_j \Phi(x_j) \quad (14)$$

$$\beta_j = C - \mu_j \quad (15)$$

4. Apply the Kuhn-Tucker conditions (KKT) complementary condition [24] by using the equality constrains from the equation (13) which result in:

$$\varepsilon_j \mu_j = 0 \quad (16)$$

$$(R^2 + \mu_j - \left\| \Phi(x_j - a) \right\|^2) \beta_j = 0 \quad (17)$$

From the equations above, the author in [24] concludes that there are three types of points. The points with $\varepsilon_j > 0$ and $\beta_j > 0$ lie outside the hypersphere in feature space and it is called Bounded Support vector or BSV. While the points with $0 < \beta_j < C$ lie on the surface and it is called Support Vector or SV. Finally the other points lie inside the sphere.

5. Use the appropriate kernel function such as Gaussian kernel [4] to represent the dot product:

$$K(x_i, x_j) = e^{-q\left\| x_i - x_j \right\|^2} \quad (18)$$

6. The distance from the sphere center to each point is defined as [24]:

$$R^2(x) = \left\| \Phi(x) - a \right\|^2 \quad (19)$$

So from equation (19) and the definition of the kernel we conclude that:

$$R^2(x) = k(x,x) - 2\sum_j \beta_j K(x_i - x_j) + \sum_{ij} \beta_i \beta_j K(x_i, x_j) \quad (20)$$

Cluster assignment

The author in [24] uses a method called complete graph (CG) to differentiate between the data point that belong to different clusters based on the following remark:

Any path that connects pairs of point that belongs to different cluster must exit from the sphere such that $R(y) > y$. To implement this idea the author [24] uses the definition of the adjacency matrix A_{ij} between pairs of points x_i and x_j:

{1 if R(y) < R 0, otherwise} *(21)*

SVC complexity

The time complexity for kernel evaluation according to the testing benchmarks for SVC algorithm proposed by [24] is $O(n^2)$ while the time complexity for the clustering labeling part is $O(N-n_{bsv})^2 n_{sv} d$ where n_{bsv} is the bounded support vector, n_{sv} is the number of support vectors, and d is the dimensionality. So the overall complexity is $O(n^2 d)$.

III. SUPPORT VECTOR CLUSTERING ENHANACMENT

Many techniques are proposed to improve the clustering labeling process for SCV and these methods are:

Support vector graph

This method [33] proposed as a modification in the method proposed in [24]. So instead of checking the linkage between all pairs of data, we consider a linkage only between points and support vectors. This method take $O(N-n_{bsv})^2 n_{sv} d$,where n_{bsv} the number of bounded support vector, n_{sv} is the number support vectors and *d* is the dimensionality.

Proximity graph technique

The previous SVC algorithm proposed by [24], despite its ability to deal with outliers and to make a cluster of arbitrary shape suffers from two problems in the cluster labeling process its efficiency become low when the number of support vector increase and it produces a false negative. So the author in [26] presents a new clustering assignment method based on proximity graph [27, 28]. In this technique instead of calculating the adjacency matrix coefficient *xi and xj* for each pairs of point *xi* and *xj* in the data space [24] which time complexity is $O(n^2 d)$ where d is the dimensional space, or calculate the A_{ij} only for pairs of point x_i and x_j where x_i and x_j is a *SV* [33], the A_{ij} coefficients are calculated for the point x_i and x_j where these point are linked by an edge which time complexity *O(nlogn).*

Gradient decent technique (GD)

The gradient decent method was proposed by [29] to treat the problem of the cluster labeling strategy in [24, 26]. Although the method proposed by [24] easy to implement but its time complexity is $O(n^2 d)$. Also despite the ability of the method discussed in [26] to reduce the time complexity of [24] to $O(n \log n)$ but it fails frequently in labeling the cluster correctly [29]. This method solves the problem by decomposing the data set into a small number of disjoint groups. Each group is represented by its candidate point and all the points that belong to the same cluster. The candidate points are labeled which result in labeling the whole data point with $O(n \log n)$ time complexity.

Improved support vector clustering

The previous method for SVC [24, 26, 29] suffers from two important problems which are computational cost and Poor labeling performance. The method [30] proposes a new support vector clustering method to overcome the problem that is attached in [24, 26, and 29]. This method performs a reduction strategy on the data set to extract the qualified subset of the data. The reduction strategy depends on Schrodinger equation [30]. This method present a new labeling strategy whose idea is to label the separate vector first and then label the other data based on labeled SVs.

The optimization part of this approach is $O(M^3)$ which is lower than the time taken by SVC which is $O(N^3)$. Table 1 shows that the time for SVC and iSVC in real data set. While the overall time taken by this strategy is $O(Nsv)^3 + N - Nsv$ where the time taken to decompose the eigenvalue is $O(Nsv)^3$ where *sv* is the number of support vector and the time taken to label the other data is $O(N - N_{sv})$. Table 2 compares the time taken by this approach with the other labeling techniques.

TABLE 1. TIME COMPARISON FOR SVC AND iSVC [30]

	SVC		**iSVC**	
	Size	***Time***	***Subsetsize***	***Time***
Liver	354	115.1	100	0.661
Sonar	208	3.32	60	0.093
wine	178	2.32	52	0.087
Iris	150	9.09	46	0.138
Vote	435	126.6	125	0.811
Diabetes	768	261.3	219	5.687
Ionosphere	351	55.47	104	0.507

TABLE 2.TIME COMPARISON OF LABELING APPROACHES [30]

	CG1	**SVG**	**PG**	**GD**
Liver	657	202	109	131
Vote	815	286	119	89
Ionosphere	1069	301	187	205

IV. COMPARISON

This section discusses previous studies that compare the algorithms used in this study with other algorithms based on different factors. Pawan [1, 9] Presents a comparative study that compare k-mean and Fuzzy c-mean in terms of time and space complexity. This study implemented on MATLAB and showed that the time and space complexity for HCM are $O(ncdi)$ and $O(cd)$ respectively and the time and space complexity for FCM are $O(ndc^2 i)$ and $O(nd + nc)$ respectively. Where n is the number of data point, c is the number of cluster, *i* is the number of iterations and d is the number of dimensions. Table 3, Table 4 and Table 5 show the result of this comparison.

TABLE 3.TIME COMPARISON FOR FCM AND HCM [1]

Number of Cluster	FCM Time Complexity	HCM Time Complexity
1	3000	3000
2	12000	6000
3	27000	9000
4	48000	12000
20	900	8

TABLE 4.SPACE COMPARISON FOR FCM AND HCM [1]

Number of Cluster	FCM Time Complexity	HCM Space Complexity
5	450	2
10	600	4
15	700	6

TABLE 5.TIME AND SPACE FOR FCM AND HCM [9]

Algorithm	Time Complexity	Space Complexity
HCM	*k*	*cd*
FCM	$O(ndc^2i)$	*O(nd+nc)*

Abbas [2] compares these algorithms in term of size of dataset, number of clusters, type of dataset and type of software used. Table 6 and figure 2 show the results.

TABLE 6.THE RELATIONSHIP BETWEEN NUMBER OF CLUSTER AND ALGORITM PERFORMANCE [2]

Performance		
Number of Cluster	*SOM*	*K-mean*
8	59	63
16	67	71
32	78	84
64	85	89

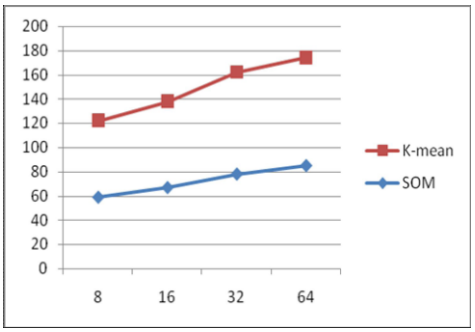

Figure2. The relation ship between number of clusters and algorithm performance

Comparison for different SVC enchantments

Many enhancements are proposed for SVC algorithm which effects on the computational complexity of this algorithm. Table 7 shows that the proximity graph and gradient decent method have the best time. But the practical total time for iSVC labeling method is the best among the other methods [30].

TABLE 7.TIME COMPLEXITY ANALYSIS FOR THE DIFFERENT SVC IMPROVMENTS

Complete graph	Support vector graph	Proximity graph	Gradient decent	Improved SVC
$O(n^2d)$	$O(n - n_{bsv})n_{sv}^2$	*O(nlogn)*	*O(nlogn)*	$O(N_{sv})^3$

CONCLUSION AND RECOMMENDATION

This study compares two important groups of clustering algorithm which are parametric and non parametric clustering algorithm. K-mean and fuzzy c-mean are a parametric clustering algorithms which require to determine the number of the cluster in a prior while SOM and SVC are a nonparametric algorithms which don't require a prior knowledge about the number of cluster and construct a cluster of arbitrary shape. The study conclude that fuzzy c-mean algorithm requires more time and space than k-mean. Also it concludes that SOM has a better performance over k-mean. Furthermore the study discusses the different enchantments for SVC such as complete graph labeling strategy, support vector graph, proximity graph, gradient decent strategy and improvement support vector clustering. The study shows that SVC is better than other clustering methods because it solves many problems which are not solved by the other clustering algorithms. As a result SVC can be applied to a wide variety of application domain. It deal with outlier and overlapping by controlling the kernel width and the softmargin constrains. Also this paper concludes that the gradient decent and the proximity graph labeling methods improve the support vector clustering time by decreasing its computational complexity, but the practical total time for iSVC labeling method is better than other methods that improve SVC.

REFERENCES

[1] Pawan, K. and Deepika, S., 2010. "Comparative analysis of FCM and HCM algorithm on iris dataset". International journal of computer application, 5(2), pp.33-37.

[2] Abbas, O., 2008."Comparison between Data clustering algorithms". The international journal of information technology, 5(3), pp.320-325.

[3] Hammouda, K.M, 2008. "A comparative study of data clustering techniques". International journal of computer science and information technology, 5(2), pp. 220-231.

[4] Ivan, G.C., Francisco A.T., de, C and M, C.P, 2004. "Comparative analysis of clustering method for gene expression time course data". Genetics and molecular biology, 27(3), pp.623-631.

[5] Hanafi, G., Abdulkader, S., 2006."Comparison of clustering algorithm for analog modulation classification". Expert Systems with Applications: An International Journal, 30(4), 642-649.

[6] satchidanandan, D. Chinmay, M. Ashish, G and Rajib, M., 2006."A comparative study of clustering algorithms". Information technology journal, 5(3), pp 551-559.

[7] Velmurugan, T and Santhanam, T., 2010."Computational complexity between k-mean and k-medoids clustering algorithm for normal and uniform distribution of data points". Journal of computer science, 6(3), pp. 363-368.

[8] Sairam, Manikandan and Sowndary., 2011. "Performance analysis of clustering algorithms in detecting outliers". International journal of computer science and information technology, 2(1), pp.486-488.

[9] pawan, K. Pankaj, V and Rakesh, S., 2010."Comparative analysis of fuzzy c mean andhard c mean algorithm". international journal of information technology and knowledge management. 2.

[10] Ujjwal, M. Sanghamitra, B., 2002."Performance evaluation of some clustering algorithms and validity indices". IEEE, 24(12), pp.1650-1654.

[11] Luai, A.S., Zyad, S and Basel, K., 2006."Data Mining: A Preprocessing Engine". Journal of Computer Science. 2 (9). pp 735-739.

[12] Han. J., Kamber. M. and et al, 2006."Data mining: Concepts and technique". 2nd edition. United state of America: Morgan Kaufmann.

[13] Jain, A.K., Murty, M.N., Flynn, P.J., 2000. "Data clustering: a review". ACM Computing Surveys (CSUR), 31(3). pp. 264-323.

[14] Anil, K.J., 2010" Data Clustering: 50 Years Beyond K-Means". Journal of Pattern Recognition Letters, 31(8).
[15] Luai, S and Zyad, S., 2006".Normalization as a Preprocessing Engine for Data Mining And the Approach of Preference Matrix", International Conference on Dependability of Computer Systems.
[16] Rajashree, D., Debahuti, M., Amiya, K.R., Milu.A., 2010."A hybridized K-means clustering approach for high dimensional dataset". International Journal of Engineering, Science and Technology. 2(2), pp. 59-66.
[17] Karthikeyani, V., Suguna, J., 2009."K-Means Clustering using Max-min Distance Measure". The 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009).
[18] Fedja, H and Tharam,S.D.,2005. "CSOM: Self-Organizing Map for Continuous Data". 3rd IEEE international Conference on Industrial Informatics (INDIN).
[19] Hsiang, C.L., Jenz, M.Y.,Wen,C.L., Tung, S.Liu., 2009."Fuzzy c-means algorithm based on PSO and mahalanobis disyance". International Journal of Innovative Computing, Information and Control. 5(12). pp. 5033–5040.
[20] kumar, P., Siri, K.W., 2010."Comparative analysis of k-mean based algorithms". International journal of computer science and network security. 10(4). pp.314-318.
[21] karthikeyani, N.V., Thangavel, K., 2009."Impact of normalization on distributed k-mean clustering". International journal of soft computing. 4(4). pp.168 - 172.
[22] Zhi, Y., Qi, Z_ and Wentai, L., 2009."Neural Signal Classification Using a Simplified Feature Set with Nonparametric Clustering". Journal Neurocomputing Journal. 73(1-3). pp.
412-422.
[23] Yiheng, C., Bing, Q ., Ting Liu., Yuanchao, L and Sheng, L., 2010."The Comparison of SOM and K-means for Text Clustering". Computer and Information Science. 3(2). pp. 268-
274.
[24] Asa, B.H., David, H., Hava T. S and Vladimir, V., 2001."Support Vector Clustering". Journal of Machine Learning Research. pp. 125-137.
[25] Cortes, C. and Vapnik, V. (1995)."Support-vector networks". Machine Learning, 20(3). pp. 273-297.
[26] Jianhua, Y., Vladimir, E.C, and Stephan, K.C., 2003."Support Vector clustering Through Proximity Graph Modeling". IEEE. 2. pp 898 – 903..
[27] Estivill, V.C and Lee, I. A., 2000."Hierarchical clustering based on spatial proximity using Delaunay diagram". In Proc. of the 9th Int. Symposium on Spatial Data Handling, pp 26 – 41.
[28] Estivill, V.C and Lee, and Murray., A. T., 2001."Criteria on proximity graphs for boundary extraction and spatial clustering". In proceedings of the 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining.
[29] Jaewook Lee and Daewon Lee., 2005."An Improved Cluster Labeling Method for Support Vector Clustering". IEEE, 27(3). pp. 461-464.
[30] Ling, P, Zhou, C.G and Zhou, X., 2010."Improved support vector clustering. Engineering Applications of Artificial Intelligence". Elsevier, 23(4). pp. 552-559..
[31] William, S.N., 2006."What is a support vector machine". Nature biotechnology. 24(12). pp. 1565-1567.
[32] Scholkopf. B., Smola. A.J., 2002."Learning with Kernels". London. MIT press.
[33] Asa, B.H, David, H and Vapnik. V., 2001."A support vector method for hierarchical clustering". MIT press.
[34] Hu, X., 2003."DB-HReduction: A data preprocessing algorithm for data mining applications". Applied Mathematics Letters.16(6). pp. 889-895.
[35] C.J. Burges.,1998. "A Tutorial on Support Vector Machines for Pattern Recognition". Data Mining and Knowledge Discovery, 2(2), pp. 121-167.
[36] Giraudel. J. L., Lek. S. 2001. "A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination". Elsevier Science. pp. 329-339.
[37] Russo. v. 2006."State of the art clustering technique Support Vector Methods and Minimum Bregman Information Principle". Master thesis.
[38] Roger. S. 2007. "Labeling clusters in an anomaly based IDS by means of clustering quality indexes". Master thesis.